UNIVERSITÀ DEGLI STUDI DI MILANO

DOTTORATO DI RICERCA IN SCIENZE DELLA TERRA

Ciclo XXXIV

DIPARTIMENTO DI SCIENZE DELLA TERRA

TESI DI DOTTORATO DI RICERCA

Statistical methods to assess rockfall susceptibility in an alpine environment: a focus on climatic forcing and geomechanical variables

GEO/05 - Geologia applicata

DOTTORANDA
Greta BAJNI
Matr. R12299

TUTOR
Prof.ssa Tiziana APUANI

CO-TUTOR
Dott. Corrado A.S. CAMERA

COORDINATORE DEL DOTTORATO
Prof.ssa Maria Iole SPALLA

Abstract

The overarching goal of the doctoral thesis was thus the development of a systematic procedure capable to examine and enhance the role of geomechanical and climatic processes in rockfall susceptibility, performed with statistically based and Machine Learning techniques. To achieve this purpose, two case studies were analysed in the Italian Alps (Valchiavenna, Lombardy Region; Mountain Communities of Mont Cervin and Mont Emilius, Aosta Valley Region).

For both case studies, Generalized Additive Models (GAM) were used for rockfall susceptibility assessment; for the Valchiavenna case study, a Random Forest (RF) model was tested too. All models were validated through k-fold cross validation routines and their performance evaluated in terms of area under the receiver operating characteristic curve (AUROC). Predictors' behaviour physical plausibility was verified through the analysis of the mathematical functions describing the predictors-susceptibility modelled relationships. Specific objectives of the two case studies differed.

The Valchiavenna case study was dedicated to testing the role of the outcrop-scale geomechanical properties in a rockfall susceptibility model. Specific objectives were: (i) the optimal selection of sampling points for the execution of geomechanical surveys to be integrated within an already available dataset; (ii) the regionalization over the study area of three geomechanical properties, namely Joint Volumetric Count (Jv), rock-mass weathering index (Wi) and rock-mass equivalent permeability (Keq); (iii) the implementation of the regionalized properties as predictors in a rockfall susceptibility model, along with the traditional morphometric variables; (iv) the investigation of prediction limitations related to inventory incompleteness; (v) the implementation of a methodology for the interpretation of predictors' behaviour in the RF model, usually considered a black box algorithm; (vi) the integration of the RF and GAM outputs to furnish a spatially distributed measure of uncertainty; (vii) the exploitation of satellite-derived ground deformation data to verify susceptibility outputs and interpret them in an environmental management perspective.

The additional geomechanical sampling points were selected by means of the Spatial Simulated Annealing technique. Once collected the necessary geomechanical data, regionalization of the geomechanical target properties was carried out by comparing different deterministic, regressive and geostatistical techniques. The most suitable technique for each property was selected and geomechanical predictors were implemented in the susceptibility models. To verify rockfall inventory completeness related effects, the GAM model was performed both on rockfall data from the official landslide Italian inventory (IFFI) and on its updating with a field-mapped rockfall dataset. Regarding the RF model, the Shapely Additive exPlanations (SHAP) were employed for the interpretation of the predictors'

behaviour. A comparison between GAM and RF related outputs was carried out to verify their coherency, as well as a quantitative integration of the resulting susceptibility maps to reduce uncertainties. Finally, the rockfall susceptibility maps were coupled with Synthetic Aperture Radar (SAR) data from 2014 to 2021: a qualitative geomorphological verification of the outputs was performed, and composite maps were produced.

The key results were: (i) geomechanical predictor maps were obtained applying an ordinary kriging for Jv and Wi (NRMSE equal to 13.7% and 14.5%, respectively) and by means of Thin Plate Splines for Keq (NRMSE= 18.5%). (ii) Jv was the most important geomechanical predictor both in the GAM (witha deviance explained of 7.5%) and in the RF model, with a rockfall susceptibility increase in correspondence of the most fractured rock masses. (iii) Wi and Keq were penalized (i.e., they had low influence on rockfall susceptibility) in the GAM model, whereas Keq showed an importance comparable to Jv in the RF model. (iv) In a complex Machine Learning model (RF), the SHAPs allowed the interpretation of predictors' behaviour, which demonstrated to be coherent with that shown in the GAM model. (v) The models including the geomechanical predictors resulted in acceptable rockfall discrimination capabilities (AUROC>0.7). (vi) The introduction of the geomechanical predictors led to a redistribution of the high-susceptibility areas in plausible geomorphological contexts, such as in correspondence of active slope deformations and structural lineaments, otherwise not revealed by the topographic predictors alone. (vii) Models built with solely the IFFI inventory, resulted in physically implausible susceptibility maps and predictor behaviour, highlighting a bias in the official inventory. (viii) The discordance in predicting rockfall susceptibility between the GAM and the RF models varied from 13% to 8% of the total study area. (ix) From the integration of InSAR data and susceptibility maps, a "SAR Integrated Susceptibility Map", and an "Intervention Priority Map" were developed as operational products potentially exploitable in environmental planning activities.

The Aosta Valley case study was dedicated to challenge the concept of "susceptibility stationarity" by including the climate component in the rockfall susceptibility model. The availability of a large historical rockfall inventory and an extensive, multi-variable meteorological dataset for the period 1990-2020 were crucial input for the analysis. Specific objectives were: (i) the identification of climate conditions related to rockfall occurrence (ii) the summary of the identified relationships in variables to be used in a susceptibility model; (iii) the optimization of a rockfall susceptibility model, including both topographic, climatic and additional snow-related predictors (from a SWE weekly gridded dataset).

Starting from an hourly meteorological dataset, climate conditions were summarized in indices related to short-term rainfall (STR), effective water inputs (EWI, including rainfall and snow melting), wet-dry cycles (WD) and freeze-thaw cycles (FT). Climate indices and rockfall occurrence time series were

paired. Critical thresholds relating rockfall occurrence to climate indices not-ordinary values (>75th percentile) were derived through a statistical analysis. As summary variables for the susceptibility analysis, the mean annual threshold exceedance frequency for each index was calculated. Model optimization consisted in stepwise modifications of the model settings in order to handle issues related to inventory bias, physical significance of climatic predictors and concurvity (i.e., predictors collinearity in GAMs). The starting point was a "blind model", i.e., a susceptibility model created without awareness of the rockfall inventory characteristics and of the physical processes potentially influencing susceptibility. To reduce the inventory bias, "visibility" masks were produced so to limit the modelling domain according to the rockfall collection procedures adopted by administrations. Thirdly, models were optimized according to the physical plausibility of climatic predictors, analysed through the smooth functions relating them to susceptibility. Finally, to reduce concurvity, a Principal Component Analysis (PCA) including climatic and snow-related predictors was carried out. Subsequently, the obtained principal components were used to replace the climatic predictors in the susceptibility model.

The key results were: (i) the 95% of the rockfalls occurred in severe (or not ordinary) conditions for at least one among the EWI, WD and FT indices; (ii) ignoring inventory bias led to excellent model performance ($0.80 \leq AUROC \leq 0.90$) but physically implausible outputs; (iii) the selection of non-rockfall points inside the "visibility mask" was a valuable approach to manage the inventory bias influence on outputs; (iv) the inclusion of climate predictors resulted in an improvement of the susceptibility model performance (AUROC up to 3%) in comparison to a topographic-based model; (v) the most important physically plausible climate predictors were EWI, WD, with a deviance explained varying from 5% to 10% each, followed by the maximum cumulated snow melting with a deviance explained varying from 3% to 5%. The effect of FT was masked by elevation. (vi) When the climate and snow related predictors were inserted in the susceptibility model as principal components, concurvity was efficiently reduced.

The inclusion of climate processes as non-stationary predictors (i.e., considering climate change) could be a valuable approach both to derive long-term rockfall susceptibility future scenarios and in combination with short-term weather forecasts to adapt susceptibility models to an early warning system for Civil Protection purpose.

Contents

Chapter 1
Preface

1.1. Introduction to rockfall phenomena

Rockfalls are classified as fast-moving landslide type, consisting of the detachment of a rock block (or several individual blocks) from a vertical or sub-vertical cliff followed by rapid down-slope motion characterized by free-falling, bouncing, rolling and sliding phases (Varnes, 1978; Selby, 1982; Cruden and Varnes, 1996). In hazard-related studies, the term rockfalls refers to volumes lower than 10^5 m^3 with negligible dynamic interaction between single blocks (Frattini et al., 2008). Due to the high energy and mobility, rockfalls are a major cause of landslide fatalities (Bunce et al. 1997; Hoek 2000; Frattini et al., 2008) and deeply affect human society and infrastructures in mountainous environments (Ravanel and Deline, 2010; Duvillard et al. 2015; Scavia et al., 2020). In the lower portion of slopes or in the valleys, rockfalls may damage roads and rail routes or properties, while at higher altitudes, these events can affect tourists and damage infrastructures such as cable cars, ski runs, trekking and climbing paths (Corò et al., 2015). In the densely frequented European Alps, public authorities are becoming aware of the increasing rockfall-related risks (Magnin et al., 2017), also in consideration of the effects of the 21st century global warming (Gobiet et al., 2014; Stoffel et al., 2014). Therefore, the development of reliable and transferable procedures to deal with such instabilities and their evolution is crucial. Several studies focusing on the analysis of rockfalls and rock mass systems were carried out at different scales: (i) at the slope scale, by means of physically-based numerical models (Gischig et al., 2011; Brideau et al., 2011; Wang and Ni, 2014; Morcioni et al., 2020); (ii) at the medium scale, including contiguous slopes, by means of ground-based monitoring systems, such as LiDAR (Gigli et al., 2014; Dunham et al., 2017; Matasci et al., 2015; 2018), or with the pioneering use of augmented reality (Zhang et al., 2019); (iii) less frequently at the regional scale, by means of hybrid physically and statistically based approaches (Frattini et al., 2008), multi-criteria decision-making methods (Cignetti et al., 2020), and rarely with statistical and machine-learning methods (Messenzehl et al., 2017; Zhou et al., 2018; Losasso and Sdao, 2018; Fanos and Pradhan, 2019; Rossi et al., 2021).

Corominas et al. (2014) analysed and distinguished conditioning (i.e., predisposing, preparatory) and triggering factors and processes leading to rockfall occurrence. Triggering factors are immediate causes acting directly, while preparatory factors are linked to a slow cumulative effect, requiring a higher

amount of time to induce a major consequence (Gunzburger et al., 2005). Nevertheless, the distinction between preparatory and triggering factors should not be considered as a dualistic concept, as they act in a continuous spatial-temporal transition (Dorren, 2003; Gunzburger et al., 2005). Also, slope systems can be described in terms of internal parameters and external factors; the former are intrinsic features of the slopes and may evolve over time due to external processes that lead to slope collapse (Volkwein et al., 2011). The internal parameters crucial for rockfall occurrence, and usually associated with a predisposing effect, are (Volkwein et al., 2011; Corominas et al., 2014):

1) Slope morphology: elevation, slope gradient, aspect, roughness and curvature concur in defining slope potential energy, meteo-climatic processes differential occurrence (e.g., wet-dry cycles, freeze-thaw cycles, permafrost), hydrological processes and local stress state distribution.
2) Geology and geomechanical properties: rock type, weathering degree and its depth, the variability of the geological structures (e.g., faults, folding), the joint intensity and discontinuity sets characteristics determine the hydro-mechanical behaviour of the rock mass system.
3) In situ stress state: a wide spectrum of topographic, tectonic, glacial loading-unloading and exhumation generated stresses interacted at several spatial and temporal scales and preferential directions, concurring in the localization of the actual rock mass system stress state (Ballantyne, 2002; Leith et al., 2014).

The main external processes responsible of rockfall occurrence, potentially acting both as preparatory and triggering processes, are:

1) Meteo-climatic processes: intense rainfall events, prolonged precipitation periods, freeze-thaw cycles and temperature fluctuations above 0°, snow dynamics (Matsuoka, 2019; Ravanel et al., 2013; Macciotta et al., 2015; Paranunzio et al., 2019; Nigrelli et al., 2018; Scavia et al 2020; Camera et al., 2021; Morcioni et al., 2022 – in press). These meteo-climatic processes are in turn responsible of degrading and weakening processes linked to modifications of the slope water circulation patterns, weathering and erosional processes, fractures nucleation and coalescence.
2) Seismicity: rock mass structure could be dramatically damaged and weakened by earthquake shaking, both at the slope scale and at the micro-scale. Large earthquakes frequently act as triggering factors for rock mass instabilities (Keefer, 1984a,b; Wasowski et al., 2011; Valagussa et al., 2014; Marzorati et al., 2002), even at consistent distances from the epicenter (up to 300 km, Stoffel et al., 2019). Nonetheless, also the long-term effects generated by earthquakes (i.e., post-seismic effect) may lead to large deformation and fracture propagation, thus representing a temporal persisting preparatory process for rock mass instabilities (He et al., 2021).

3) Human activity: anthropic activities such as quarrying and mining, excavation for infrastructures, vibration due to blasting are responsible of the alteration of rock slope geometry and stress distribution (Selby, 1982; Dorren, 2003; Hantz et al., 2003).

1.2. Objectives

The assessment of landslide susceptibility has been largely addressed in the literature, at several geographical scales and in various environments (Reichenbach et al., 2018). However, researchers have been traditionally focused on shallow landslides susceptibility. Conversely, rockfalls are commonly investigated through physically-based models, dealing with runout analyses rather than on the spatial prediction of potentially critical rockfall sources. Therefore, the successful development of susceptibility models for an underexplored landslide type is an open research question. Thus, the overarching goal of the doctoral thesis was the development of a systematic procedure capable to promote and analyse the role of geomechanical and climatic processes in rockfall susceptibility, performed with statistically based and Machine Learning techniques, and to contribute to define rockfall susceptibility mapping procedures in the context of mitigation strategies.

The thesis general objectives were:

(i) To select appropriate geomechanical and climatic processes and synthesize them in spatially distributed predictors for susceptibility modelling of rockfall occurrence in Alpine environments.

(ii) To ascertain the suitability and the (in)completeness of the available input data - i.e., rockfall inventories, geomechanical and meteorological datasets - and to propose strategies to address the related uncertainties.

(iii) To assess the function of the newly introduced predictors during the susceptibility modelling phase, not only through traditional quantitative performance metrics, but also in terms of plausibility and coherency of their physical-geological behaviour within the model.

(iv) To determine the consequences of the selection either of different susceptibility modelling algorithms or different model configuration setups on the outputs.

(v) To move towards non-stationary landslide susceptibility models by including climate-related predictors.

To explore these issues, two distinguished case studies were set up and explored; they were chosen for their peculiarities and data availability, and for their potential to satisfy distinguished requirements of

the proposed research questions. The two study areas are: (i) the Valchiavenna territory, located in Lombardy Region in the Italian Central Alps and (ii) the Mountain Communities of Mont Cervin and Mont Emilius, located in Aosta Valley Region in the Italian Western Alps. Anthropic and seismic factors were not accounted in the analyses, as not considered of primary importance in the selected study areas.

The principal specific objective for the Valchiavenna case study was to test the role of outcrop-scale geomechanical properties in a rockfall susceptibility model. The project benefitted from the availability of an extensive geomechanical and geomorphological dataset for the area, developed during many years of surveys carried out by the Geoengineering Group of the University of Milan. Three target geomechanical properties, obtainable from the processing of geomechanical field survey data, were chosen as potential rockfall susceptibility predictors, namely Joint Volumetric Count (Jv), rock-mass weathering index (Wi) and rock-mass equivalent permeability (Keq). As statistically-based and Machine Learning models require spatially distributed variables, the regionalization of the selected geomechanical properties was necessary. The already available geomechanical dataset was not suitable for this purpose, as the available geomechanical surveys were clustered along roads and thus not enough homogeneously distributed. Therefore, the first specific objective for the Valchiavenna case study was the optimal selection of additional locations for geomechanical surveys execution and data collection, by means of the Spatial Simulated Annealing technique. After an intensive field campaign, for the regionalization, different deterministic, regressive and geostatistical techniques were tested and compared in term of quantitative performance and physical reliability. Consequently, the obtained regionalized geomechanical properties were implemented as predictors in a rockfall susceptibility model along with the traditional morphometric variables, performed both using Generalized Additive Models (GAM) and Random Forest (RF). Both models allowed the analyses of predictors' behaviour, by means of Component Smoothing Functions (CSF) and Shapely Additive exPlanations (SHAP), respectively.

The available rockfall inventory for Valchiavenna was the result of the integration of the official landslide Italian inventory (IFFI) and a geomorphological field-mapped rockfall dataset. Three GAM models were produced and compared: a topographic model, a model containing both topographic and geomechanical predictors, and a third model with the further addition of the geological component. The aim was to examine the variation of the susceptibility spatial patterns amongst models, by introducing knowledge about the geomechanical conditions and geological features of rock masses. To assess potential implications of inventory incompleteness, a complete GAM model including all the predictors was performed also on the solely IFFI dataset. Regarding RF, one model was produced, including all the morphometric, geomechanical and geological predictors, and SHAP values were used for outputs

interpretation. The GAM and RF outputs were then compared to each other to examine the coherency of predictors behaviour and geomorphological plausibility of the susceptibility maps. The resulting maps were integrated in order to quantify accordance-discordance areas between the two models in predicting susceptibility and therefore to furnish a spatially distributed measure of uncertainty. In addition, rockfall susceptibility maps were integrated with satellite-derived ground deformations data (Synthetic Aperture Radar-SAR), both qualitatively, for a geomorphological verification of the outputs, and quantitatively, producing operative combined maps potentially tailored for environmental management and planning purposes.

The principal specific objective for the Aosta Valley case study was to define a robust procedure to include a climate component in the rockfall susceptibility model. In the dynamic framework of climate change, this part of the research project was aimed at representing a step towards undermining the limiting concept of a time-invariant susceptibility. The fundamental idea was to couple the concepts of susceptibility spatially-distributed predictors and of Intensity-Duration thresholds. This could represent an essential and propaedeutic procedure to subsequently update susceptibility maps with data coming from future climate projections. The set-up of this part of the project benefitted from the availability of a large historical rockfall inventory and an extensive, multi-variable meteorological dataset for the referecnce period 1990-2020 (i.e., approximately a three-decade reference period allowing the calculation of Climate Normals as defined by the World Meteorological Organization, WMO 1989, 2007). At the beginning of the PhD project, the meteorological dataset available was in the form of station-based temperature and precipitation data series. During March 2021, the "Centro Funzionale Valle d'Aosta" made also available a grid-based meteorological dataset of temperature and precipitation.

The first part of the study addressed the identification of climate processes related to rockfall occurrence. An exploratory analysis between meteorological conditions recorded in several time- frames before rockfall occurrence and the average, or ordinary, conditions recorded at the station in the reference period was carried out for this scope. The idea was to verify if the majority of the rockfalls (more than 50%) occurred in severe, or not ordinary, meteorological conditions related to different processes (climate index value for the specific process larger than the 75^{th} percentile of its distribution), thus reasonably linking rockfalls and climate through a cause-effect relationship. The processes analysed were rainfall (sub-daily data), effective water inputs (daily data, both considering rainfall and snow melting), wet-dry episodes and freeze-thaw cycles. For each climatic process recognized having a role in rockfall occurrence, empirical critical thresholds were defined. For each climatic index, specific procedures were set up in order to define, during the reference period, the number of independent

threshold exceedance events. Consequently, mean annual threshold exceedance frequencies for each climate index were used as synthetic climatic predictor for rockfall susceptibility modelling. These climatic predictors were produced both from the station-based hourly dataset (with consequent regionalization) and from the grid-based hourly dataset. The two sets of variables were alternatively used in the models.

The second part of the study was aimed at carrying out a rockfall susceptibility model by means of GAMs, including both topographic, climatic and additional snow-related predictors obtained from a Snow Water Equivalent (SWE) weekly gridded dataset (from Camera et al., 2021). A four-step procedure was carried out in order to manage the crucial issues of inventory bias, physical significance of climatic predictors and concurvity (i.e., predictors collinearity in GAMs) by stepwise modifications and improvements of the model setup. The first step dealt with the creation of a "blind model" i.e., a susceptibility model created purposely ignoring the rockfall inventory characteristics and the physical plausibility of the functions relating climatic predictors to susceptibility values. The second step focused on the implementation of a strategy to reduce the inventory bias effects on susceptibility outputs. To achieve this, the model domain was reduced according to the actual reporting activity of rockfall events by regional Forest Corps, through the creation of proper "visibility" masks. The third step concentrated on optimizing models according to the physical plausibility of the mathematical functions describing the behaviour of climatic predictors in defining susceptibility. The fourth step implemented a Principal Component Analysis including climatic and snow-related predictors. Subsequently, the obtained principal components were used to replace the climatic predictors in the susceptibility model, in order to minimize concurvity effects. Outputs from the four steps were compared and discussed in terms of plausibility, susceptibility spatial patterns and quantitative performance.

1.3. Thesis structure

Chapter 1 has defined the investigated instability phenomena and has highlighted the general purpose and the rationale of the research, as well as given a brief overview of the methodological approach. A glossary for the acronyms used in the thesis will be also provided in Section 1.4 to facilitate readability. The thesis is distinguished in three major chapters. Chapter 2 is focused on the theoretical concepts of geo-spatial modelling, fundamental for the purposes of the project. It deals with sampling optimization strategies, regionalization techniques and a detailed summary of the state-of-the-art on landslide susceptibility modelling.

Chapter 3 and Chapter 4 are dedicated to the Valchiavenna and Aosta Valley case studies, respectively. It was deemed appropriate to offer a comprehensive and specific literature review at the beginning of each of these two chapters, in order to introduce the reader to the research questions and objectives addressed. In particular, Chapter 3 deepens the literature related to rock mass instability processes from a geomechanical point of view, while Chapter 4 addresses the topic of meteo-climatic processes as predisposing and triggering factors. In both chapters, these introduction sections are named "Specific objectives and research questions". The same applies for the Section "Study area" and the final Section named "Discussion and future perspectives".

The core of Chapter 3 is linearly organized in the sections: "Data", "Methods" and "Results". The first topic addressed is the acquisition and calculation of geomechanical target properties and their regionalization. Secondly, rockfall susceptibility models by means of Generalized Additive Models and Random Forest are carried out. Then, the outputs of the two distinct modelling algorithms are examined and discussed separately. Finally, strategies are tested for the integration of the susceptibility maps coming from the two modelling techniques and coupled with ground deformation satellite data.

Due to procedure complexity, Chapter 4 is structured in three main parts. "Results" of each part directly follow the corresponding "Methods" section. The first part deals with the definition of climatic processes influencing rockfall occurrence and thresholds definition. The second part provides the procedure for the realization of the climatic predictors. The third part deals with rockfall susceptibility modelling.

A final and overall concluding remark is given in Chapter 5, "Conclusions".

1.4. Glossary

The acronyms are presented in alphabetic order.

A-DInSAR techniques: Advanced Differential Interferometric Synthetic Aperture Radar techniques
AIC: Akaine Information Criterion
ARPA: Agenzia Regionale per la Protezione Ambientale (Regional Agency for Environmental Protection)
ATLR: Altitudinal Temperature Lapse Rate (°C)
AUROC: Area Under the ROC Curve
Cm: degree day factor (mm/degree-day C)
CSF: Component Smoothing Function
CV: Cross Validation
DEM: Digital Elevation Model
DS: Distributed Scatter (for SAR)
DSGSD: Deep Seated Gravitational Slope Deformation

ECDF: Empirical Cumulative Distribution Function
edf: effective degrees of freedom
EWI: Effective Water Inputs (mm)
EWIind: EWI independent
EWIper: EWI persistence
FT: Freeze Thaw Cycles
FTN: Freeze Thaw Normal
GAM: Generalized Additive Models
GCV: Generalized Cross-Validation
GIS: Geographic Information System
GLM: Generalized Linear Models
GPS: Global Positioning System
GSI: Geological Strength Index
GWR: Geographically Weighted Regression
Idp: Inverse Distance Weighting power
IDW: Inverse Distance Weighting
IFFI: Inventario dei Fenomeni Franosi in Italia (Italian Landslide Inventory)
ISRM: International Society of Rock Mechanics
JRC: Joint Roughness Coefficient (-)
Jv: Joint Volumetric Count fractures/m^3)
KED: Kriging with External Drift
Keq: Equivalent Permeability (m/s)
LiDAR: Light Detection And Ranging
LOO-CV: Leave-one-out Cross Validation
LOS: Line of Sight
MAP: Mean Annual Precipitation (mm/year)
mDD%: mean Decrease in Deviance Explained
MESS: Multivariate Environmental Similarity Surface
ML: Maximum Likelihood
MLR: Multiple Linear Regression
MR: Melt Rate (mm/day)
mtry: the number of candidate variables randomly selected at each split during trees growing in RF
NDVI: Normalized Difference Vegetation Index
nlos, hlos and elos: the direction cosines of the LOS
NMRSE: Normalized Mean Root Square Error
nsCV: non-spatial Cross Validation
ntree: number of trees in RF
OK: Ordinary Kriging
OOBE: Out-Of-Bag Error
PCA: Principal Component Analysis
PS: Permanent Scatter (for SAR)
RDN: Rainy Days Normal
REML: Restricted Maximum Likelihood

RF: Random Forest
RMR: Rock Mass Rating
ROC: Receiver Operating Characteristics Curve
SAR: Synthetic Aperture Radar
sCV: spatial Cross Validation
SHAP: Shapely Additive exPlanation
SSA: Spatial Simulated Annealing
STR: Short Term Rainfall (mm)
SWE: Snow Water Equivalent (m)
SWEep: average number of melting events occurring over 16-day periods in a hydrological year
SWEmax: maximum amount of melting recorded over 32-day periods in the whole data series
SWI: SAGA Topographic Wetness Index
***TEF_a*:** annual threshold exceedance frequency
TPS: Thin Plate Slpine
UNIMI: Università degli Studi di Milano
Vlos average annual velocity of PS and DS in the satellite LOS direction
Vslope: Vlos projected in the direction of the steepest slope
WD: Wet and Dry Cycles
Wi: Weathering index (-)
WMO: World Meteorological Organization

Chapter 2
Fundamentals of geospatial modelling and analysis

This chapter outlines the theoretical background of the most relevant concepts and techniques of spatial analysis and modelling used throughout the thesis. Three main topics are explored: (i) sampling design, used for optimizing field work and geomechanical data collection for the Valchiavenna case study; (ii) regionalization techniques, employed for the creation of spatial distributed landslide susceptibility predictors; (iii) susceptibility modelling, concerning the background and the state of the art on landslide susceptibility by statistical and machine learning techniques, from landslide inventory types and related issues to model validation and uncertainties evaluation. This last section deals with landslide susceptibility in general, without focusing on a peculiar landslide type, as the modelling framework is generalizable. A more specific introduction with a comprehensive literature review on rockfall predisposing and triggering processes will be given at the beginning of the following Chapter 3 and Chapter 4, dealing with geomechanical and climatic processes relevant for rock mass instability, respectively. This choice was deemed appropriate for sake of linearity, and to offer the reader an immediate connection between the literature research questions and the variety of analyses performed within the two different case studies.

2.1. Sampling optimization strategies

Wang et al. (2012) defined sampling as the "selection of a subset of individuals from within a population to estimate characteristics of the whole population"; it is a fundamental step to plan efficiently field surveys and to project monitoring networks in a multitude of environmental and geological applications. The optimization of sampling locations has the goal of overcoming both the costly, time-consuming and sometimes redundant intense (or exhaustive) sampling and the sparse sampling, which, although economically advantageous, could miss important features or areas (Cochran, 1977; Wang et al., 2013).

De Gruijter et al. (2006) offered a comprehensive overview of sampling techniques, which could be broadly subdivided in convenience sampling, purposive sampling and probability or random sampling. Convenience sampling is when sampling is limited to specific locations controlled by their accessibility (e.g., along the road network) and it is considered arbitrary. Although convenient from both an economic

and time perspectives, it could be affected by biases. Purposive sampling (or non-probability sampling) selects the sampling locations by satisfying a specific purpose; this could be made in a subjective manner, dependent on the surveyor sensibility and experience, or by optimizing an objective function related to the purpose of sampling. Probability sampling selects sampling locations at random; the probability of selecting a specific location is assumed to be known and could be useful for successive statistical inference from the data. This sampling mode is also referred as design-based approach.

Two well-known design-based sampling techniques are Simple Random Sampling and Stratified Random Sampling. For the former, a pre-specified number of sample locations is randomly selected among geographical coordinates of the area, with equal and independent probability of selection; the latter is done by applying Simple Random Sampling to sub-regions (i.e., strata), within which the probability of sample selection could differ (de Gruijter and Brus, 1997).

Purposive, non-probability sampling techniques have gained increasing popularity in recent years. If the goal is to use sampled data to estimate values at unobserved locations (i.e., regionalization), then it is possible to optimize the non-probability sampling in various ways (Brus, 2019). Several non-probability sampling designs exist for mapping purposes:

(i) Regular grid and spatial coverage sampling, based on the optimization of geometrical rules and distances between samples. Although simple and straightforward methodologies, they are entirely based on the spatial coordinates of the locations and on the number of affordable sampling points (Brus, 2019). The method is considered not suitable to deal with the non-uniform, highly variable, mountainous environment.

(ii) K-means sampling and conditioned Latin Hypercube sampling (Minasny and McBratney, 2006), based on the optimization of the representativeness of the samples in the covariates space. In an extremely variable environment as the alpine valleys, a very high number of sampling points may be required to be representative of the different combinations in the covariate space. In detail, k-means sampling tends to concentrate the sampling points where, in the multivariate distribution, the density of points is largest, i.e., in the most “common” environments (de Gruijter et al., 2010). In Conditioned Latin Hypercube a series of intervals (marginal strata) is defined for each covariate; the interval breaks are chosen such that the numbers of pixels in the marginal strata are equal. Subsequently the algorithm selects the sampling points by minimizing a criterion which is a function of the number of sampling points in the marginal strata and the correlation matrix of the environmental features (Ma et al., 2020). For this reason, this method is particularly

suitable when the variable of interest could be represented by a complex (i.e., not linear) combination of numerous covariates which, in turn, need to be available as spatially distributed variables.

(iii) Model-based sampling, which assumes the availability of a model of spatial variability to optimize grid spacing or coordinates of sampling locations. The main drawback is that this sampling design approach must rely on a quite consistent number of previous collected data, or on the availability of enough resources to carry out a reconnaissance sample survey. In other words, the availability of a pre-existing sample is necessary to build the spatial model to develop the final sampling design (Brus, 2019). These methods consist in the minimization of functions related to the spatial structure of the pre-existing sample (mainly the kriging variance from Ordinary Kriging or Kriging with External drift in case of covariates inclusion). Model-based sampling usually leads to the selection of locations spread out throughout the study area (this particularly useful if the previous data were clustered) and without necessarily assuming covariates (Brus, 2019).

A single best sampling design technique cannot be determined; the best technique is site-specific as it depends on the method that will be adopted for the subsequent regionalization of the property of interest (Brus, 2019). When covariate maps related to the variable of interest are not available, the best sampling method are the geometry related ones and model-based sampling. Conversely, when covariates related to the variable of interest are available, all the three methods presented above may be used, depending on the expected relationship type (i.e., linear or not) between the variable and covariates. However, usually, the regionalization method at the stage of the sample design is not already decided. For this reason, the covariate inclusion at the sampling stage could be a very sensitive issue as it is necessary to decide in advance which covariates explain part of the variation of the variable of interest. When the covariates are used in the sampling optimization but not used for the subsequent regionalization, the sampling could be suboptimal and *vice versa* (Brus, 2019).

The above discussed sampling methods were evaluated; the goal was to select the most suitable for the optimization of data collection and regionalization of geomechanical properties in the Valchiavenna area. Considering (i) the complex mountainous environment characterizing the area, (ii) the numerosity of already available geomechanical data and their clustering along roads, (iii) the complexity of the properties to be regionalized, not necessarily involved in a priori recognizable relationship with readily available morphometric covariates, it was deemed appropriate to focus on model-based sampling techniques. Amongst this category, Spatial Simulated Annealing - SSA (Van

Groenigen and Stein, 1998; Van Groenigen et al., 1999, 2000) was adopted for the present study and is briefly presented hereafter. This technique is able to deal with kriging variance minimization. It was widely and successfully used in a multitude of environmental, hydrological, and agronomic applications (Chen B. et al., 2013; Barca et al., 2015; Wang et al., 2016; Scudiero et al., 2016; Li et al., 2021). Spatial Simulated Annealing is an iterative random search procedure (Brus, 2019) that optimizes a custom target function based on a model of spatial variability built on pre-existing points. Randomness is associated to the selection of the candidate sampling configurations: a sequence of sampling configurations is generated, where each new proposed configuration is obtained by slightly modifying the current one, by a perturbation consisting of a transformation of the sampling locations over a vector with random length and a random direction (Van Groeningen et al., 1999). The SSA algorithm implements an iterative procedure as at each step *i*, it considers a neighbouring configuration *Si* of the current configuration *S0* and probabilistically decides whether to move to *Si* or stay in the previous configuration. The transition between the current configuration to the new configuration is controlled by an acceptance probability *Pt* defined by the Metropolis Criterion (Van Groeningen et al., 1999), which depends on the value of ϕ, i.e., the so-called fitness function to be minimised (e.g., the kriging variance associated to the pre-existing points) for the two configurations *S0* and *Si*, and on a parameter called Annealing Temperature *T* (Barca et al., 2015):

$$P_t(S_0 \rightarrow S_i) = 1 \qquad \text{when} \qquad \varphi(S_i) < \varphi(S_0)$$
$$P_t(S_0 \rightarrow S_i) = \exp\left(\frac{\varphi(S_0) - \varphi(S_i)}{T}\right) \qquad \text{when} \qquad \varphi(S_i) > \varphi(S_0) \qquad \text{Eq. 2.1}$$

The larger *T*, the larger the probability that a new proposed sample with an increase of the fitness function instead of a further minimisation is accepted. The advantage of using the SSA method is that even combinations with worse configurations than *S0* are accepted with a certain probability, meaning that suboptimal configurations (i.e., local minima of the fitness function) are discarded (Brus and Heuvelink, 2007); indeed, as *T* is gradually decreased during the procedure, the acceptance probability of worse samples (i.e., only local minima) gradually moves towards zero (Brus, 2019). Initial temperature *T0* is one of the most sensitive parameters to avoid local minima (Ameur, 2004) and for SSA performance (Barca et al., 2015) and it is usually set as suggested by Triki et al. (2005) through the equation:

$$T0 = -\frac{\overline{\Delta\varphi^{(+)}}}{\ln(X_0)} \qquad \text{Eq. 2.2}$$

which represents the ratio between the average of all the rejected fitness function values and the probability of acceptance of worse configurations. The Annealing Temperature decreases during the iterative process according to a given cooling scheme. The geometric cooling scheme has demonstrated the best balance between algorithm running time and the rate of convergence to the global optimum (Nourani and Andresen 1998):

$$T_{i+j} = \alpha_{cool} T_j \qquad \text{Eq. 2.3}$$

where j is the current temperature index and, $\alpha cool$ is the cooling rate, usually belongs to the range 0.800−0.995 (Barca et al., 2015).

Beyond the sampling approach selected, a useful tool to verify the representativeness of the obtained sampling scheme could be the Multivariate Environmental Similarity Surface (MESS; Elith et al., 2010), which is widely used especially in the ecology and biology fields (e.g., Owens et al., 2013; Reygondeau et al., 2017). The higher the associated MESS value, the more common the environment of the points is, while negative values indicate a novel environment, meaning one or more covariates are not fully represented by the survey locations (Camera et al., 2017a). Moreover, in such cases that mapping methods involving covariates would be used, MESS allows to evaluate the extent and location of the extrapolated areas.

2.2. Regionalization techniques

The optimization of a spatial sampling has in general two main purposes in the environmental field: designing an efficient monitoring network and planning field (or remote) surveys to obtain valuable point-based datasets, necessary to estimate spatially distributed variables. Spatially distributed data (or spatially continuous surfaces) play a significant role in environmental sciences and management (Li et al., 2011).

The process of estimating a target variable at unsampled locations is usually referred as spatial interpolation (Li et al., 2011) or prediction (Veronesi et al., 2019). The former has a stricter definition as it is the process of estimating the values of a variable at points within the same region, more often (but not necessarily) in the value range of the observations at the sample locations. Also, interpolation considers distances and mutual positions between sampled and unsampled points and usually refers to the geographical space. Prediction has a broader meaning as the estimation could rely on a more general "feature" space, including time. Moreover, prediction at points outside the region (or beyond value ranges of the feature space covered by sampled points) could be also more specifically referred as

extrapolation (Burrough and McDonnell, 1998). As the majority of environmental variables could be considered as spatially dependent random variables (Oliver et al., 1989), explainable through the Regionalized Variable Theory (Matheron, 1963), regionalization is used too as a synonym of interpolation. Throughout the whole thesis, interpolation and regionalization could be intended as synonyms, while prediction will be used as the preferred term when dealing with susceptibility modelling, benefitting from its broader connotation.

Li and Heap (2014) enumerated the principal features distinguishing and characterizing regionalization methods, which are summarized below:

- Global/local methods: the former use all available data in the study area to estimate a general trend, while the latter operate on short-range variations.
- Exactness: some methods generate estimates, which exactly match the observed values, while the others are considered inexact methods.
- Deterministic/stochastic: while stochastic methods provide both estimation (the deterministic part) and uncertainties (the stochastic part), deterministic methods only produce estimations.
- Abrupt/Gradual: this characteristic distinguishes between methods producing discrete surfaces (i.e. maps) and smoothing surfaces.
- Convex/non-Convex: Convex methods estimates are always valued between the values range of the observed values, whereas non-convex methods can yield estimates outside of the range of the observed values.
- Univariate/Multivariate: methods using only samples of the target variable in deriving the estimation are termed univariate methods, whereas methods that also use explanatory variables are referred to as multivariate.

For the evaluation of the interpolation performance, the establishment of an appropriate confidence level, or of an acceptable error, is necessary when dealing with environmental data, as they are often involved in decision-making and land management. The modelling (interpolation) function usually leads to an error that incorporates two quantities: its bias and its variance (James et al., 2013). On the one hand, bias refers to the approximation error of a function that can accurately fit only data that follow a strict pattern and is linked to the ability of the function to adapt its shape. On the other hand, functions more adaptable to the pattern of data could result in higher variance (Veronesi et al., 2019).

Bennet et al. (2013) offered a very detailed categorization of quantitative performance approaches. They divided the approaches in:

(i) direct statistical values comparison of the whole dataset (e.g., comparing real and interpolated data mean, range, skewness etc.).

(ii) Coupling real world and modelled values, via concurrent comparison (e.g., hit rate, false alarm ratio etc.), key residual methods (e.g., Root Mean Square Error, RMSE, or Mean Absolute Error, MAE).

(iii) Testing the ability of the model in preserving data patterns (e.g., through Pearson's correlation coefficient between modelled and measured data).

(iv) Indirect metrics, useful to detect possible model overfitting (e.g., Akaike Information Criterion, AIC).

(v) Data transformation into different domains to highlight aspects of a model's behaviour not revealed in the original domain.

To obtain error and performance metrics, sample data should be split in subsets in order to both develop the model and to validate it. The most frequently used method for this subdivision is cross-validation (CV), which includes three sub-categories (Kohavi, 1995):

(i) Hold-out CV, where the original dataset is split into two groups, one for model development and one for model evaluation. A popular split percentage between the two is 80/20 or 70/30.

(ii) K-fold CV, where the original dataset is split in k sets, one used for model evaluation and the remaining k-1 for development. The procedure is repeated k times, so that each fold is used k-1 times for development and once for model performance evaluation.

(iii) Leave-one-out CV (LOO-CV), where n-1 data points are used for model development and only one point is used for validation (e.g., a k-fold CV where n=k). The procedure is repeated for all data points. In general, k-fold and LOO-CV are less affected by size and position of group splitting than hold out CV.

Nevertheless, it should be highlighted that quantitative metrics should not be the only criteria to define the best performing model or to state whether a model is useful. A qualitative assessment done by experts may become essential in highly complex or data-scarce situations, where a model behaviour that resembles the real system trends is more useful than quantitative metrics (Bennett et al., 2013).

Many interpolation methods are available in literature, including deterministic, regressive, and geostatistical methods, with different level of complexity in terms of interpretation and model construction (Attorre et al., 2007). Even sampled data density and topographic complexity are important in finding the best interpolation method (Li and Heap, 2014). A brief overview on the main characteristics of the most used interpolation methods is given in the next paragraphs. Inverse Distance Weighting (IDW), Thin Plate Splines (TPS), Multiple Linear Regression (MLR), Geographically Weighted Regression (GWR), Kriging (in particular Ordinary Kriging OK and Kriging with external Drift KED) will be presented.

IDW This technique estimates the values of a variable at unsampled points using a linear combination of values at sampled points, weighted by an inverse function of the distance from the point of interest to the sampled points. The weights are expressed as (Shepard, 1968):

$$W_i = \frac{1}{d(x, x_i)^p} \qquad \text{Eq. 2.4}$$

where x is the point where the estimate is wanted, *xi* is the i^{th} point where observations are available, d is the distance between the two locations, and p is the power parameter that controls the influence of remote observation on the estimation. The assumption is that sampled points closer to the unsampled point of interest have more similar values to it than those further away; weights decrease as the distance increases, especially when p is set to high values, resulting in a more local spatial interpolation (Isaaks and Srivastava, 1989). The most popular choice of p is 2 (Li and Heap, 2008).

TPS Splines consist of polynomials describing locally a piece of a line or surface (i.e., they are fitted to a small number of data points exactly) and fitted together so that they join smoothly (Burrough and McDonnell, 1998; Webster and Oliver, 2001; Li and Heap, 2008). Thin plate splines were developed by Wahba and Wendelberger (1980) for climatic data. Specifically, the smoothing parameter of TPS is calculated by minimising the generalised cross validation function - GCV (Hutchinson, 1995) and are based on the minimisation of curvature and enforcement of smoothness (Li and Heap, 2008).

MLR Multiple Linear Regression is not a spatial interpolator in the technical sense; however, it could be used to relate multiple covariates to the variable of interest through a mathematical function, which is eventually used to estimate the value of the latter in unsampled locations. The approach is identical to simple least squares regression with a dependent and an independent variable; when the covariates are multiple, the workspace is more complicated than a simple two axes Cartesian plane. The standard MLR model for spatial estimation is based implicitly upon the assumption of spatial stationarity in the relationship between the dependent variable and the covariates, and the estimated parameters are assumed to be constant over space.

GWR Geographically weighted regression (Fotheringham et al., 2002) is based on the traditional regression framework but incorporates local relationships; differently from MLR is a regional regression method that can be used to investigate the non-stationary relationship between the dependent and explanatory variables, thus accounting for spatial heterogeneous processes.

The equation for a typical GWR would be:

$$y_i(u) = \beta_{0i}(u) + \beta_{1i}(u)x_{1i} + \beta_{2i}(u)x_{2i} + \cdots \beta_{mi}(u)x_{mi} \qquad \text{Eq. 2.5}$$

where the notation $\beta_{0i}(u)$ indicates that the parameter describes a relationship around location u, thus being specific to that location. Weights are conditioned on the location u and hence change for each location. The weighting scheme is also known as kernel; several kernel shapes are possible, but the most used are the gaussian and bilinear types. The kernel is a function of the distance between the target point and the n observations around that point, falling in a distance range called bandwidth. As the bandwidth gets larger the weights approach 1 and the local GWR model approaches the global MLR model. The combination of geographically weighted estimators, kernel and bandwidth can be referred to as a local model. In terms of sensitivity, the bandwidth is more sensitive than the kernel shape and could be selected as fixed or adaptive. The former indicates that the bandwidth is the same for each local estimation, the latter allow the kernel to vary according to possible sampling irregularities by increasing and decreasing the bandwidth size depending on sample data density throughout the study area.

GEOSTATISTICS and KRIGING Geostatistics was originated from the work in geology and mining by Krige (1951) and it was successively formalized by Matheron (1963) with his theory of regionalised variables. The key concept of geostatistics is "When a variable is distributed in space, it is said to be regionalized, […] geostatistical theory is based on the observation that the variabilities of all regionalized variables have a particular structure" (Journel and Huijbregts, 1978).

Kriging is the definition of a family of generalised least-squares regression algorithms used in geostatistics, based on the variogram model, which provides an estimation of the autocorrelation pattern in the study area (Webster and Oliver, 2007). It can be classified as a BLUE (best linear unbiased estimation) methodology, as in every point the estimation of the variable of interest is given by a linear combination of the weighted neighbouring observations and with the sum of weights equal to one (Hofstra et al. 2008). The variogram is created by averaging the semi-variances, which are calculated as follows:

$$\gamma'(x_i, x_0) = \gamma(h) = \frac{1}{2} var[Z(x_i) - Z(x_0)] \qquad \text{Eq. 2.6}$$

where Z is the regionalized variable of interest, h is the distance between points x_i and x_0 and $\gamma(h)$ is the semi-variogram.

A plot of semi-variances against *h* is known as the experimental variogram, which is characterized by three essential elements (Li and Heap, 2008): (i) the nugget, a positive value of semivariance at distances close to 0, which reflects the variance due to sampling errors or the spatial variance at distances shorter than the minimum sample spacing; (ii) the range, which is a value of distance at which an asymptotic semi-variance value is reached; (iii) the sill, which is the asymptotic value on the semi-variance axis reached at the range. The samples separated by a distance larger than the range are spatially de-correlated or independent. If the ratio between sill and nugget is close to 1, then most of the variability related to the variable of interest is non-spatial (Hartkamp et al., 1999).

Variogram modelling is the procedure involving the approximation of the experimental variogram with a mathematical model (i.e., variogram model); the choice of the variogram model is extremely important for the spatial interpolation. Several variogram models are available, e.g. Exponential, Spherical, Gaussian, Power and the nested sum of one or more simple models (Pebesma, 2004; Webster and Oliver, 2001). The sampling design is extremely important for kriging as, to obtain a robust representation of the autocorrelation structure, it requires samples to be separated by a variable range of distances (lags), so that the variogram can capture the spatial complexity of the area from short to large separation distances (Li and Heap, 2008). Environmental properties show commonly different autocorrelation structures in different directions, thus sometimes, anisotropic semi-variogram modelling is required; the most commonly employed model for anisotropy is the geometric anisotropy, meaning the semi-variogram reaching the same sill in all directions and with the same variogram model, but at different ranges. In such cases, a maximum and minimum ranges are modelled, revealing the maximum and minimum correlation directions, respectively. A focus on two different and widely used kriging

estimators, namely Ordinary Kriging (OK) and Kriging with External Drift (KED), is briefly presented in the following paragraphs.

Ordinary Kriging assumes that the local means are unknown but constant within a search area, thus limiting the stationarity domain to a local neighbourhood; moreover, the local means are not necessarily closely related to the population mean. OK is based on the spatial correlation structure of the data to determine the weighting values, and the spatial correlation between data points determines the estimated value at unsampled locations. It also assumes a normal distribution of the data points and can account for local fluctuations of the mean.

Kriging with External Drift incorporates the local trend within the neighbourhood search window as a linear function of a smoothly varying secondary variable instead of as a function of the only spatial coordinates as done by OK (Goovaerts, 1997). The assumption is that the trend of the primary variable must be linearly related to that of the secondary variable, which must be known both at all sampled points and at all points which need to be estimated Pebesma (2004).

2.3. Landslide susceptibility modelling: state of the art

2.3.1. Overview and definition

Landslide susceptibility is the likelihood of a landslide occurring in an area on the basis of local environmental conditions (Brabb, 1984); it is the process of predicting "where" landslides are likely to occur (Guzzetti et al., 2005, 2006; Reichenbach et al., 2018), by giving a measure of the degree to which a terrain can be affected by future slope movements. In mathematical terms, susceptibility can be defined as a probability of spatial occurrence of slope failures, given a set of geo-environmental conditions and independently from any size characteristic of the landslide itself (Guzzetti et al., 2005). For this reason, susceptibility is different from landslide hazard, which could be defined as the probability that a landslide of a given magnitude will occur in a given period and in a given area, in relationship with a defined trigger (Guzzetti et al., 2005).

Methods for landslide susceptibility could be subdivided in qualitative or quantitative (Reichenbach et al., 2018); the former evaluate susceptibility heuristically, using descriptive terms, while the latter produce numerical outputs, i.e., estimates of the occurrence likelihood of landslide phenomena. More specifically, these methods can be grouped into the following main categories (Guzzetti, 2005, Reichenbach et al., 2018; Lombardo et al., 2020):

(i) Geomorphological mapping, which strongly depends on the experience of the investigators in determining the sliding potential of a slope and it is essentially impractical over very large areas. Moreover, it does not provide quantitative estimates hampering its employment for hazard assessment.

(ii) Analysis of landslide inventories, which is obtained preparing landslide density maps and inherently depends on the quality and completeness of the available landslide database.

(iii) Heuristic or index-based methods, which are again dependent on the subjective choices of the investigators in weighting the known instability factors in causing landslides.

(iv) Deterministic, physically- based models, which exploit the existing mechanical laws controlling slope instability. Although providing numerical measures, these models require several input parameters and a large dataset regarding mechanical and hydrological properties of the slope materials, which may not be available or difficult and expensive to acquire, especially over large and heterogeneous territories.

(v) Statistical predictions models, which exploit the functional relationships between a set of instability factors and the past and present distribution of landslides. Practically a binary classification model is fitted to a spatial data set containing information on the presence and absence of past landslides (response) and some associated preparatory environmental factors (predictors). Finally, the resulting classification rule, which allowed to identify conditions that promoted or favoured past instability, is applied to all spatial units containing information on these environmental conditions (Steger et al., 2016a). The entire process is also referred as *susceptibility modelling*.

The next sections will be dedicated to the deepening of the statistical models for landslide susceptibility, focusing on several aspects and challenges. Among them, the selection of the mapping unit, the landslide inventory (dependent variable), the environmental predictors (independent variables), the modelling framework (which operates as the functional relationship) will be discussed. In addition, a final section regarding the evaluation of model performance, uncertainties and plausibility of the results of the process will be introduced.

2.3.2. Mapping units

The preliminary requirement for landslide susceptibility modelling involves the selection of an appropriate mapping unit. As defined by Guzzetti (2006), it is a geographical domain characterized by the maximization of internal homogeneity and the between-unit heterogeneity in terms of a set of ground

conditions. Reichenbach et al. (2018) categorized and discussed advantages and disadvantages of different mapping units used in the field of landslide susceptibility.

The different categories include (i) grid cells ("pixels"), (ii) unique condition units, and (iii) slope units. The selection between them is dependent on scale, quality and type of geo-environmental predictors, landslide size and type, modelling technique used and scope of landslide susceptibility assessment (Guzzetti, 2006).

Grid cells are the most popular among landslide susceptibility modellers, thanks to their simplicity and applicability at all resolutions and scales. Reichenbach et al. (2018) however listed some drawbacks related to their usage. Particularly concerning for the authors is the possibility of a pixel to be physically representative of a landslide process. This is particularly true when dealing with morphometric and geometrical predictors, unless the grid-cell size is small compared to the landslide size. In that case, very-fine grid cells are suitable for modelling small landslide, as they allow to capture their morphological signature in detail but may result geomorphological insignificant for large or deep-seated landslides. A similar problem concerns the final susceptibility zonation, which may be difficult to interpret and use operationally, whether several artefacts related to the pixels are present. When this situation occurs, the authors suggest defining clear and unambiguous criteria for post-processing to improve results readability.

Unique condition units (Bonham-Carter, 1994) are obtained by intersecting all the geo-environmental layers considered important for susceptibility modelling and are simply obtainable in a GIS environment. The main problem with unique conditions is that each continuous predictor need to be categorized in a small number of classes. Nevertheless, class selection introduces a heuristic and subjective component in the process, and in other words, an unquantifiable uncertainty. Moreover, even small digitalization imprecisions could result in poorly significant units.

Slope units are hydrological terrain units bounded by drainage lines (Carrara, 1983; Carrara et al., 1991, 1995; Guzzetti et al.,1999), corresponding to a slope in the geomorphological point of view, being physically appropriate for the scope (i.e., landslide susceptibility modelling). The main advantage of slope units is that their shape and size may be tailored to the landslide type and size: a slope unit may correspond to an individual slope, an ensemble of adjacent slopes, or a small catchment. In particular, they can provide representative statistics of large, deep-seated landslides. Recently, Alvioli et al. (2016) proposed a specific software for the automatic delineation of slope units, overcoming their principal disadvantage, related to the difficulties in their manual identification. A possible limitation could be that

drainage-based subdivision may not reflect some peculiar geomorphological and geological characteristics; thus, additional post-processing should be made. Furthermore, a non-optimal construction of slope units would lead to the proxy variable characterization into the slope-units as too smoothed (Amato et al., 2019; Baet al., 2018; Rotigliano et al., 2011, 2012; Martinello et al., 2020).

Some authors focused on the comparison between a grid-cell and a slope unit approach, both in terms of quantitative performance and readability of the results. Van Den Eeckhaut et al. (2009) combined the susceptibility zonation obtained by applying both a pixel-based and slope unit approach by means of a heuristic procedure, stating that this procedure helped in increasing the interpretability of the susceptibility zonation. Martinello et al. (2020) carried out a pixel-based landslide susceptibility model and zoned the scores into 10 different types of slope units, obtained by differently combining two half-basin and four landform classification coverages. The authors found out that the predictive performance of this approach was slightly lower than the pixel-based model, balanced with an increasing readability of the final map. Jacobs et al. (2020) found out that a slope unit-based approach outperformed the pixel-based one, when dealing with a point-based inventory with unknown accurate location within or in vicinity of the landslide. In the work of Ba et al. (2018), the slope unit approach slightly outperformed the grid-cell one. However, the authors pointed out that for the slope unit-based model, the same susceptibility level is assigned to the whole unit, making it difficult to determine within which part of the slope landslides tend to occur, an important information for next step analysis, e.g., run-out modelling, hazard and risk assessment.

2.3.3. Landslide inventory

The dependent or response variable is typically used in a binary structure, expressing the presence or absence of landslides in each mapping unit used to partition a study area (Lombardo et al., 2020). The landslide presence can be obtained from landslide inventories and catalogues (Guzzetti et al., 2012; Van Den Eeckhaut and Hervás, 2012). The indispensable information is the location of occurrence of the events. Additional information can consist in the date of occurrence, the landslide type, the failure mechanism, the causal factors, involved volumes and damage caused by mass movements that have left recognizable traces in an area (Guzzetti et al., 2000; Van Westen, 2008). Landslide inventory is considered the most important component in the susceptibility process (Van Westen, 2008). Different techniques and tools are available for the preparation of landslide inventory maps (Guzzetti, 2006, 2012); their selection depends on the purpose of the inventory, the extent of the study area, the scale of the base maps, the characteristics of the available imagery as well as the investigators skills and the resources available (Guzzetti et al., 2000; Van Westen et al., 2006; Guzzetti, 2012). Tools and techniques for the

collection of landslide information may be categorized in conventional and innovative methods (Van Westen, 2008; Guzzetti, 2012; Reichenbach et al., 2018).

Conventional methods include field mapping, including Mobile GIS and GPS for attribute data collection, and visual interpretation of stereo aerial photographs. The former has objective difficulties relying on the visibility and discernability of landslide phenomena from the investigator viewpoint (Santangelo et al., 2010), the availability of trained and experienced geomorphologists and the possible obliteration of landslide characteristics due to human activities, e.g. forestation, agriculture (Guzzetti, 2012; Petschko et al., 2013, 2014). Conversely, aerial photographs interpretation, even if considered a benchmark in landslide mapping, may be limited by the lack of interpretation standards (i.e. the identification is based on the investigator sensibility and experience), and by the availability of adequate coverage and time-series length.

Innovative techniques comprehend analyses of high and very-high resolution digital elevation models (acquired by means of airborne laser profilers and LiDAR sensors). The analyses are performed both by means of image analysis of single or multiple acquisition and through (semi) automated classification by exploiting different terrain characteristics as surface curvature variation and slope threshold. This is typically done on a pixel basis (Chen G. et al. 2012; Hussain et al. 2013), which however often results in imprecise outcomes for complex morphologies. An object-oriented classification procedure can be alternatively adopted (Drăguţ and Blaschke 2006; Lu et al. 2011; Stumpf and Kerle 2011; Drăguţ and Eisank 2012; Hölbling et al. 2015, 2017), which is capable of exploiting embedded and scaled geomorphological features characterizing landslide phenomena (Kurtz et al., 2014). Otherwise, a combination of remote sensing and morphometric analysis is preferable (Mondini et al., 2011; Ciampalini et al. 2016; Du L. et al. 2019). A second category of innovative techniques comprehend the exploitation of satellite derived products, both from passive (optical) and active (radar) sensors. The visual interpretation of optical images and secondary products (e.g. panchromatic, composite, false colour, pan-sharpened) is a valuable alternative to aerial photography and could be used in (semi) automated processes (Martha et al., 2012; Van Den Eeckhaut et al., 2012; Hölbling et al. 2015, 2017; Catani, 2021). Recently, Google Earth™ imagery has been employed as a source of information too (Conoscenti et al., 2016; Broeckx et al., 2017). Another approach is to concentrate on satellite multispectral information to construct derivative maps (e.g., Normalized Difference Vegetation Index, NDVI). These maps are therefore used in combination with aerial or optical products to visually detect landslides, or as inputs for (semi) automatic classification by means of e.g., index thresholding, clustering, change detection and object-oriented image analysis (Liu et al., 2002; Hervás et al., 2003;

Cheng et al., 2004; Rosin and Hervás, 2005; Borghuis et al., 2007; Yang and Chen, 2010; Martha et al., 2011; Parker et al., 2011; Scaioni et al., 2014; Behling et al., 2014; Hölbling et al. 2015; Heleno et al., 2016). Furthermore, the application of interferometric techniques to radar images has proved to be another powerful tool for landslide detection, especially at the large scale (Strozzi et al., 2006; Colesanti and Wasosku, 2006; Meisina et al., 2013; Agostini et al., 2014; Ciampalini et al., 2016; Raspini et al., 2017; Rosi et al., 2018). It consists mainly in the exploitation of A-DInSAR techniques - e.g., PSInSAR (Ferretti et al., 2000), SqueeSAR (Ferretti et al., 2011), Small Baseline Subset (SBAS) (Berardino et al., 2003). More specifically, radar images are useful to individuate landslide typology and to update existing inventories with landslide state of activity (Antonelli et al., 2019).

Based on the type of mapping, landslide inventory maps can be classified as archive or geomorphological inventories (Guzzetti, 2012; Reichenbach et al., 2018). Archive inventories are obtained by e.g. newspaper, road maintenance companies, fire-brigade, administrations, interviews (Van Westen, 2008) and "Citizen Science" (Juang et al., 2019). Geomorphological inventories can be further classified as historical, event-related, seasonal or multi-temporal (Guzzetti, 2012). An historical inventory records several landslides over a period of tens, hundreds or thousands of years (Galli et al., 2008), independently from a particular triggering event. On the contrary, an event-related inventory reports landslide caused by a single trigger (e.g. seismic event or rainfall event). Multitemporal (years or decades) or seasonal inventories are instead prepared on the basis of multiple aerial images or satellite products. Geomorphological inventories may suffer of incompleteness due to the easy and rapid obliteration of shapes and geomorphological features due to erosional and vegetational processes, reactivations, or human interference (Reichenbach et al., 2018). However, whilst the inventory is representative of the abundance and distribution of landslide phenomena and the incompleteness is rather balanced and not biased, it should not be problematic in terms of landslide modelling (Petschko et al., 2013, 2014, 2016) as consistency is more important than completeness (Reichenbach et al., 2018). Event-related inventories are less appropriate for susceptibility model construction, as not depending only on geo-environmental conditions but even on the triggering event characteristics, thus their generalization should be carried out with caution, as specific predictors representing the pattern and extent of the meteorological (i.e., a particular rainfall or snowmelt event) or seismic (i.e., the Peak Ground Acceleration – PGA of a given earthquake) triggers are included in the analysis and influence the landslide spatial occurrence (Guzzetti et al., 2012; Reichenbach et al., 2018). However, these types of inventories resulted to be appropriate for susceptibility model validation (Rossi et al., 2010) and are successfully implemented as supporting tool for early-warning operative system (e.g., Segoni et al., 2015; 2018; Bordoni et al., 2020). Seasonal or multi-temporal inventories are not surprisingly considered

the optimal datasets for susceptibility modelling (Galli et al., 2008); these inventories are specifically suitable for testing the long-term performances of susceptibility models and defining their learning curve (Guzzetti et al., 2004; 2005) along with the possibility to test dynamic environmental predictors (if the period covered by the inventory is long enough). Nevertheless, as they are time and resource consuming to prepare, they are quite rare and of limited extent (Galli et al., 2008; Guzzetti et al., 2012).

The quality of a landslide inventory depends on its accuracy, which in turn is conditioned by the completeness of the map and the geographical and thematic correctness of the information recorded (Galli et al., 2008; Guzzetti et al., 2012). Subjectivity, experience, measuring errors or imprecision related to the landslide inventory preparation are all sources of parametric uncertainty (i.e., a component of the epistemic uncertainty, which is the uncertainty related to the missing knowledge of the phenomena), which propagate in the subsequent modelling and analysis (Petschko et al., 2014). The completeness level of an inventory is unknown (Malamud et al., 2004); however, if incompleteness and errors are not random but systematic, they could induce a bias in the inventory, thus introducing systematic modelling errors (Petschko et al., 2014; Steger et al., 2016a; 2017).

Recently, some researchers focused on the quantification of issues as landslide inventory positional accuracy and completeness on the modelling results (Petschko et al., 2013, 2016; Steger et al. 2016a, 2017). Steger et al. (2016a) highlighted that the occurrence of positional errors in mapping landslide inventories affected modelled relationships and variable importance assessments. Furthermore, they found that the propagation of positional error in the modelling process was not only linked to the size of the landslides under consideration and to their spatial representation (e.g., points at the landslide scarp or inside the body, polygons), but also to the modelling characteristic of the territory (i.e., raster resolution and mapping unit), to the study area morphology and to the complexity of the applied model. To deal with these issues the authors suggested, in such situation in which a mapping update of the landslide inventory is not feasible, to generalize input data to a coarser scale and to opt for simple and easily interpretable modelling algorithms.

As regarding positional accuracy, a previous work of Petschko et al. (2013) compared modelling landslide susceptibility by representing presence either as single point for the main scarp or as point randomly selected in a landslide polygon. The authors observed very small differences in the predictive performance and final aspect of the susceptibility maps. Successively, Petschko et al. (2016) compared the susceptibility modelling results using as inventory both a representative set of landslide polygons and a substantially complete inventory of main scarps related points. The trade-off between temporal

requirements for compilation and performance was found to be essentially better for the point related inventory. Ozturk et al. (2020) explored the effect of selecting landslide information from the available inventory in different positions of the landslide body, finding out that sampling the landslide presence points at the landslide toe (i.e. deposit) improved the model performance by 10% over a model using for training the entire landslide polygons. Hussin et al. (2016) tested different landslide representations (i.e., scarp centroids, points populating the scarp and entire scarp polygons) obtaining model highest performance when sampling shallow landslides as grid points and debris flow scarps as polygons.

Incompleteness is the other frequent drawback that may affect landslide inventories. For instance, in some contexts, landslides in forested areas are overrepresented than in agricultural areas, where landslide features are more frequently masked by human activities (Bell et al. 2012; Petschko et al. 2016; Conoscenti et al. 2016). Conversely, analysis of multi-temporal images may be positively biased towards unforested areas, as dense vegetation could mask landslide morphologies (Jacobs et al., 2016). Another very common situation that favours systematic incompleteness is related to inventories coming from administrations or public reports, which usually overrepresent landslide closer to infrastructures and roads (Guzzetti et al., 1999; Steger et al., 2016b; Bajni et al., 2021b). Steger et al. (2017), by synthetically creating biased inventories of different degrees (i.e., reducing a substantial complete inventory by pre-selected percentages), figured out that highly biased inventories produced very high validation results but distorted relationships with geo-environmental predictors and geomorphological implausible landslide susceptibility maps. More specifically, they relate the high performance to the presence of bias-describing predictors (e.g., land cover, distance from roads or elevation). The authors discouraged the exclusion from the model of the bias-related predictors, as a cascading effect of misleading relationships between landslide presence and such predictors (named confounding predictors) could occur. As an alternative approach they proposed the application of mixed-effect logistic regression models, which resulted in an attenuation of the influence of bias-describing predictors in the predictions, by modelling them as random effects instead of fixed effects (as usually geo-environmental predictors are modelled).

Another approach to deal with inventory bias related to a systematically unbalanced survey of landslides along roads and infrastructures was offered by Bornaextea et al. (2018) and Knevels et al., (2020). These authors created an effective surveyed area, delineated by an automatic procedure that accounted for the actual visibility of slopes during the geomorphological field mapping phase, starting from the available GPS tracks. They trained the statistical model inside the effective surveyed area and eventually applied the resulting model to the entire study area. A similar approach was adopted by Meinhardt et al. (2014),

who used the viewshed GIS algorithm from roads to select the area in which to train the susceptibility model. More recently, Steger et al. (2021) suggested to change perspective and to simultaneously predict landslide susceptibility effects and a landslide intervention index (i.e., areas where damage-causing and infrastructure-threatening landslides are likely to occur). In other words, they changed the modelling subject, from landslide susceptibility to areas affected by damaging landslides, exploiting in a positive and useful way the inherent bias in the inventory.

As landslide susceptibility model requires landslide absence points along with landslide locations, their selection is equally important; however, they are not readily available and need to be reasonably generated. Conoscenti et al. (2016) compared two different strategies: the extraction from randomly selected circles having a diameter equal to the mean width of the landslide source areas and randomly distributed points, which is the most traditional technique used in literature. Their results highlighted that the former approach is preferable in terms of predictive performance. Zhu et al. (2018), concerned about the lack of uniform standards in selecting an exclusion mask or buffer around landslide source to select absence points, compared two presence-absence methods and two only-presence methods. Their findings, in accordance with the previous studies of Zaniewski et al. (2002) and Engler et al. (2004), showed that the presence-absence methods constrain the over-prediction of susceptibility values, concluding that absence data are necessary in susceptibility modelling. Zhao et al. (2020) compared the modelling results deriving from a presence-only, presence-absence and pseudo-absence approaches. The latter approach, which resulted to be the best performing, does not assume that a landslide cannot occur at the selected sites, but only provides a large sample (in this case 25 times higher than the presence points) representing conditions available in the region. Zhu et al. (2019) used different thresholds in terms of environmental dissimilarity between the absence and presence data to quantify the reliability of candidate negative points. In other words, they separated presence and absence data in the feature space rather than on a geographical basis. They proposed a Similarity Based Sampling and tested its application by means of three different statistical models, concluding that the best performance was achieved with a dissimilarity threshold of 0.5 (in a range between 0 and 0.9). Hong et al. (2019) compared different methods, i.e., the random selection in the geographic space and four methods based on the feature (or environmental) space, to recognize an “eligible area” in which to select absence points. They observed that, in terms of performance, the size of absence points needed to be increased when the eligible area was reduced (due to the application of the different selection methods). More specifically a 1:1 ratio was optimal when the eligible area was about the 99% of the total study area, a 1:5 ratio was optimal when the eligible area was between 50 and 70% of the study area, and a 1:100 ratio when the eligible area was around 30% of the study area.

2.3.4. Geo-environmental predictors

In addition to the response variable or dependent variable, represented by landslide and non-landslide information, landslide susceptibility models require a set of geo-environmental predictors, representing the independent or explanatory variables (Guzzetti et al., 1999). In their review, Reichenbach et al. (2018) classified and quantified the frequency of use of different geo-environmental variables in landslide susceptibility models and found out that terrain morphology related variables are the most common (25.5%), followed by geological variables (19.1%), land cover (17%) and hydrological variables (17.6%). Van Westen et al. (2008) grouped predictors into two classes, static and dynamic (i.e., based on their need to be updated regularly or not). As an example, morphology and geology belong to static predictors, while land use and hydro-climatic factors are dynamic predictors.

Variables related to morphometric aspects of slopes and topography are unquestionably effective in predicting landslide likelihood (Fabbri et al., 2003; Van Westen et al., 2008). These variables basically come from the post-processing of a DEM. While elevation and slope are directly linked to slope instabilities, which are processes mainly controlled by gravity, other morphology-related variables such as aspect, curvature and terrain roughness have a less immediate and more local link with landslide predisposition (Reichenbach et al., 2018). As these topography related predictors strongly depend on the quality and resolution of the DEM (Tarolli, 2014), landslide susceptibility modelling is benefitting from the increasing availability of high-resolution DEM (Reichenbach et al., 2018). However, a finer DEM resolution does not necessarily imply either a more accurate and performing susceptibility model, or the optimalization of the derived geo-environmental predictors (Chen L. et al., 2019; Chen Z. et al., 2020; Rabby et al., 2020a). Other more complex terrain attributes (e.g., flow direction, drainage density, flow accumulation, Topographic Wetness Index) may be derived from the available DEM and they are mainly introduced as hydrological and topo-climatic processes proxies relevant for slope stability such as snow accumulation and duration, incoming solar radiation variability, drainage efficiency, local moisture patterns, sediment storage, permafrost distribution and probability, efficiency of weathering activity, plant colonization (Van Westen, 2008; Messenzehl et al., 2017 and references therein).

Geology related predictors are often included as they are classified in the basic geological maps available for the area (i.e., along with their unit or formation regional names) leading to misleading or unclear relationship with landslide susceptibility and making it difficult to compare the effect of geological features in different and distant areas (Reichenbach et al., 2018). A more appropriate representation of geology for landslide susceptibility should deal with the geotechnical behaviour (e.g., strength, cohesion) and properties (e.g., thickness, weathering, grain size, texture, permeability etc.) of the

investigated materials rather than with its chronostratigraphic subdivision, thus a reclassification in terms of physical or mechanical properties is advisable (Reichenbach et al., 2018; Van Westen, 2008). Segoni et al. (2020) performed a sensitivity analysis on the use of the geological information in various ways for landslide susceptibility modelling. They performed different tests either using lithologic, or chronologic, structural, paleogeographic, and genetic units. Even if the genetic units approach (i.e., categorizing geology as magmatic rocks, metamorphic rocks, clastic rocks, organogenic rocks and soils) performed better than the other subdivision for their case study (including the most commonly used lithologic subdivision), the most significant outcome of their work was that a comprehensive representation of geology, combining the different geological parameters together, resulted in the best performing model, as it captured the multifaceted connotation in a complex study area.

Regarding structural data, the most used related predictors are bedding attitude (Clerici et al., 2002; Coe and Harp, 2007; Ruff and Czurda, 2008; Santangelo et al., 2015; Messenzehl et al., 2016) and distance from faults (Hong et al., 2016; Du et al., 2017; Yi et al., 2020; Rabby et al., 2020b). On the one hand, a problem with structural and bedding data is that they are time consuming to collect in the field and challenging to interpolate for landslide susceptibility assessment at the regional scale (Reichenbach et al., 2018). On the other hand, distance from faults predictor suffers from several problems as it is often based on "subjective" distance classes based on the modeller sensitivity, and seldom considers the state of activity and size (i.e., the extent of the influence of the faults on the surrounding rock masses in terms of mechanical weakening) of the fault-zones.

An interesting use of the presence of old deep-seated landslides as predictor for secondary landslides was attempted by Carrara et al. (1991). The authors based the inclusion of this predictor on the basis of a possible alteration of mechanical properties of the rock masses located in the areas affected by the more ancient deep-seated deformations.

Land cover related predictors are widely used in landslide susceptibility (e.g., Glade, 2003; Knevels et al., 2020; Reichenbach et al., 2014; Reichenbach et al., 2018; Bordoni et al., 2020; Camera et al., 2021) and comprehend vegetational and land use characteristics and could be obtained both from aerial photographs and through the automatic and semi-automatic classification of optical and, more recently, multispectral satellite imagery. Land cover conditions slope stability and some land cover types could be more prone to instability. However, the relationship is not always unambiguous and generalizable; for instance, Carrara et al. (1991) found out that forested area favoured stability but in an adjacent study area Carrara et al. (1995) discovered that forested areas were more prone to landslides. The presence

and abundance of vegetation may be also represented by a numeric continuous predictor rather than a land use class, such as Normalized Difference Vegetation Index (NDVI), easily obtainable from optical satellite imagery.

Reichenbach et al. (2018) listed two additional "distance to linear features" predictors, which are widely used in the related literature and expressed concern about their inclusion in a susceptibility model. The first is distance to river, with the aim to capture the destabilizing effect of river incision favouring slope instability. This could be quite reliable for hydrologically controlled landslides or otherwise for rockfall activity, if the incision may lead to shape overhanging slope morphologies at the valley bottom. However, rarely authors limited the distance calculation to the actual slope, ignoring the presence of the divides or use this predictor only in hydrologically related contexts. The second predictor is represented by distance to roads, with the aim of capturing the disturbance of road cuttings on the natural slopes; Reichenbach et al. (2018) concluded that the distance metric is not able to capture these disturbance effects, which may be more local and geotechnical related. Moreover, as already discussed in the previous section, distance to roads could also be a bias-descripting predictors, limiting its meaningfulness in the susceptibility process.

Recently some authors raised attention on several issues regarding geo-environmental predictors for landslide susceptibility modelling, which are seldom addressed by the related literature. Reichenbach et al. (2018) argued that authors seem to be more interested in testing increasingly complex modelling algorithms and model ensembles rather than in convincingly discussing the geomorphological and physical significance and relevance of the single geo-environmental variables in the analyses. This issue is exacerbated in such cases where predictors deriving from very different spatial scales are used together in the susceptibility modelling process; indeed, the predictor physical meaning at a given scale should be consistent and relevant to the landslide type and size. In this regard, Steger et al. (2016b, 2021) extensively focused their work on the discrepancies between high quantitative model performances and implausible predictors behaviour. Camera et al. (2021) introduced climate variables within a shallow landslide susceptibility analysis, which, even with a slight increase of model performance, resulted to be physically plausible and consistent with the investigated phenomena.

Another limitation to the traditionally approach to landslide susceptibility modelling is that predictors are usually considered as static and, as a result, susceptibility is considered stationary (Lombardo et al., 2020). The assumption of stationarity does not hold especially in mountainous environments, where the occurrence of landslides is deeply connected with climate-related processes such as intense rainfall,

snowmelt, freeze-thaw cycles (Luino et al., 2020; Lucas et al., 2020; Stumvoll et al., 2020; Subramanian et al., 2020; Schilirò et al., 2021; Bajni et al., 2021b), which are expected to vary in the future due to global warming (Beniston et al., 2018; Gobiet et al., 2014). Nevertheless, studies including climate related variables are quite rare, with only 2.8% of them including a precipitation-related predictor and only 0.3% including other climatic predictors (Reichenbach et al., 2018). Different approaches were tested to include rainfall-related variables in landslide susceptibility modelling: (i) precipitation is included in the form of mean annual rainfall, mean monthly rainfall and rainy days frequency (e.g., Broeckx et al., 2018; Chen and Li, 2020; Fang et al., 2020; Nahayo et al., 2019; Nhu et al., 2020); (ii) precipitation is included in event-based susceptibility studies, where a particular intense rainfall event is summarized in variables representing multiple day-maximum cumulated precipitation and used to model post-event rainfall induced landslides (e.g. Kim et al., 2015; Knevels et al., 2020); (iii) stationary variables based susceptibility models are updated for early warning purposes by coupling them with thresholds exceedance (Segoni et al., 2015, 2018) or with additional temporal statistical models (Bordoni et al., 2020). However, all these approaches have some inherent limitations as the weather conditions leading to landslide occurrence are widely variable in terms of amount, duration and intensity (Perruccacci et al., 2017). Including precipitation as an average predictor, again fail in capturing the non-stationary connotation of landslide occurrence. The event-based approach is valid only for a limited time frame and a single precipitation event characteristic. The early warning approach, even if undoubtedly effective for civil protection purposes, do not formally include the climate-related variable in the modelling process. In this regard, the recent work of Camera et al. (2021) in Aosta Valley (Italy) tried to give an answer both to the problems related to the inclusion of a rainfall-related predictor and to the gap related to the inclusion of other climate-related variables. Authors investigated the relationships between landslide occurrences and intense rainfall and snowmelt events (period 1991–2020). Successively, they set up a susceptibility model including the effective annual number of rainfall events with intensity–duration characteristics above a defined threshold and the average number of melting events occurring in a hydrological year. These two variables together accounted for 5% of the explained deviance and their introduction led not only to a slight increase in the model performance but made the model adaptable to future climate change projections (i.e. they introduced non-stationarity in the process). In recent years, the inclusion of climate change scenarios in landslide susceptibility is becoming an appealing, and undeniably essential, challenge for researchers, but still few investigators have attempted to address it (Fan et al., 2013; Kim et al., 2015; Gassner et al., 2015; Shou and Yang, 2015).

Non-stationarity of landslide susceptibility may be also modelled in response to land cover or land use changes. Meusburger and Alewell (2009) accounted for the temporal change of land cover by comparing two different susceptibility maps from 1959 and 2000, with different land cover information. They observed poor model performances of the 1959 model on the landslide inventory from 2000, concluding that the temporal change of environmental factors chiefly influences the entire modelling process. Reichenbach et al. (2014) found that a significant susceptibility spatial distribution variation was the result of the decrease of the extent of bare soils in Sicily (Italy). Persichillo et al. (2016) elaborated different scenarios to assess the influence of land use changes on shallow landslide susceptibility modelling in three areas in NW Italy. In particular, they tested both land use modifications related to natural evolution of the landscape and to human activity (e.g. cultivated land abandonment). The latter was found to influence the increase of high susceptibility areas. Chen W. et al. (2019) ended up in similar findings in SE China. Indeed, land use change from natural to human colonized lands was responsible of increasing landslide susceptibility in the study area. Authors suggested that land use planning is vital to tackle this increasing abundance of potentially unstable areas.

Samia et al. (2017a, 2017b, 2018, 2020), Temme et al., (2020) and Lombardo et al. (2018, 2020) recently raised an additional issue regarding landslide susceptibility modelling. Indeed, they discussed the missing ability of traditional predictors to consider the spatial relationship among landslide occurrences in different mapping units. More specifically, adjacent, neighbouring, and distant landslide are considered equally by the models. Samia et al. (2017a, 2017b, 2018, 2020) in a study area in Umbria (Italy) found that occurring landslides may attract future landslides, as the latter tend to occur inside or in the immediate vicinity of the previous ones. They called this effect “path-dependency”, which was inserted in their several works both as a spatial dependency and a spatio-temporal dependency. They introduced these effects as predictors in a landslide susceptibility model, as a function of the spatio-temporal distance of earlier and nearby landslides, by means of a space–time clustering (STC) measure derived from Ripley’s space–time K function implemented on the available point-based multi-temporal landslide inventory. The characteristic timescale of this effect was about 17 years, and the characteristic spatial scale was about 60 m and was characterized by an exponential decay, observing a substantial improvement of model performance. Based on their findings, Temme et al. (2020) formalized a new nomenclature for landslide occurrence: (i) uncorrelated landsliding, when landslides are common but do not imply a correlation with geo-environmental variables; (ii) correlated landslding, when landslides are common and have correlation with geo-environmental variables; (iii) path-dependent landsliding to describe such situations where there is a correlation with previous landslides. This last category was further classified in “reactivation” or “continuation” when the same material is involved in future

landslides, “local activation” when a previous instability causes local changes leading to additional instabilities, “remote activation” if a landslide causes changes in the landscape (not necessarily nearby) that cause a later landslide. Lombardo et al. (2020) proposed a novel approach for spatio-temporal prediction of landslides exploiting a Log-Gaussian Cox Process (LGCP), which considers each landslide in a population as an individual realization from a continuous-space process, i.e., landslides are the result of a stochastic point process driven by an unknown intensity function. In other words, they exploited the spatio-temporal clustering of the landslide process to consider the spatial, the temporal, and the spatio-temporal landslide latent effects among adjacent terrain mapping units, same mapping units but subsequent time intervals, and both conditions together, respectively. For latent effects, authors mean effects not captured by the traditional geo-environmental predictors, observing that explicitly introducing this latent component significantly improved the model performance. The approach was also previously tested and applied successfully to model populations of rainfall–induced (Lombardo et al., 2018, 2019b) and seismically–triggered (Lombardo et al., 2019a) landslides in terms for spatial predictions only. However, as stated by the authors themselves, this approach needs a detailed and accurate multi-temporal inventory and, at least when introducing the temporal component of the latent effect, it introduces a borderline definition between landslide susceptibility and hazard, suggesting a strong influence of time on the spatial occurrence of landslides.

2.3.5. Models

The functional relationship describing the interplay between the landslide inventory and the geo-environmental predictors is analysed through a susceptibility model (Carrara, 1983; Chung and Fabbri, 2003; Guzzetti et al., 2006). Several statistical and machine learning models are available for this scope, each with both advantages and shortcomings (Guzzetti et al., 1999, Brenning, 2005; Glade and Crozier, 2005). They can be broadly grouped in (Reichenbach et al., 2018; Merghadi et al., 2020): (i) Statistical Methods s.s. such as Logistic Regression, which is the most popular (e.g., Steger et al., 2016b, 2017; Rossi and Reichenbach, 2016; Lin, 2017; Cama, 2016; Wang, 2013; Lombardo and Mai, 2018), Generalized Additive Models (e.g., Brenning, 2008, Jia et al., 2008, Park and Chi, 2008; Goetz, 2011, 2015; Muenchow, 2012; Petschko, 2012, 2014; Bordoni et al., 2020; Camera et al., 2021) and Weights of Evidence (e.g., Sterlacchini et al., 2011; Thiery et al., 2007; Vakhshoori and Zare 2016; Roy et al, 2019; Kouli et al., 2014; Neuhäuser et al, 2012); (ii) Machine Learning methods such as Support Vector Machines and Tree-based Ensembles (e.g., Micheletti, 2014; Goetz et al., 2015; Chang et al., 2019; Dou et al., 2020; Merghadi et al., 2020); (iv) Neural Networks (e.g. Ermini et al., 2005; Falaschi et al., 2009; Wang et al., 2019; Caniani et al., 2008;Gomez and Kavzoglu, 2005); (v) Multicriteria decision analysis (e.g. Castellanos Abella and Van Westen, 2008; Akgun, 2012; Feizizadeh and Blaschke, 2013;

Feizizadeh et al., 2014;Lorentz et al., 2016). In multicriteria decision analysis a subjective component in weighting factors may be present, Machine Learning algorithms provide high performance but often tend to overfit, while linear models are not able to capture possible non-linearities in the functional relationship between response and explanatory variables, as they have limited flexibility (Brenning, 2005; Goetz et al., 2011; Petschko et al., 2014). Reichenbach et al. (2018) observed a trend in the research studies testing increasingly complex methods; however higher complexity does not guarantee better or more sound results and rather may lead in misleading interpretations if done by non-expert users.

Another recent common practice is to apply several modelling frameworks to the same dataset and compare or combine the results. Model comparison is chiefly carried out in terms of model quantitative performance (e.g., Pham et al., 2019, 2020; Althuwaynee et al., 2014; Abedini, 2019; Chen, 2019, 2020; Pourghasemi and Rahmati, 2018) and less frequently by comparing more pragmatical metrics such as model interpretability (Goetz et al., 2015) and geomorphological plausibility of the output map (Steger et al., 2016b). The availability of different models and, consequently, different susceptibility zonation, could hamper the effective and practical application of the models in land use planning and management; indeed often, the qualitative performance is similar between models (Sterlacchini et al., 2011) and thus not sufficient to establish the "best" model for an area (Huabin et al., 2005; Chacón et al., 2006; Reichenbach et al., 2018). Conversely, more fruitful results may be obtained by combining models in different ways rather than just comparing them, in order to offer an optimal version of susceptibility zonation and both to reduce and quantitively describe uncertainties (e.g., Rossi et al., 2010; Di Napoli et al., 2020; Chen, 2018; Youssef et al., 2015, Choi et al., 2012, Andan et al., 2020; Rossi and Reichenbach, 2016).

In reality, there is not a "correct" model in a suite of competing models (Elith et al., 2002) and the choice need to be guided by the specific objectives of the model along with quantitative measures. Goetz et al. (2015) effectively summarized the possible criteria leading to an optimal model selection. Firstly, a good quantitative performance is undoubtedly a necessary condition, which is influenced in turn by the algorithm feature selection procedures, modelling parameter selection, sampling design and pre-processing of predictors. Secondly, the interpretability of the model in terms of predictors behaviour is an advisable requirement (Brenning et al., 2012b), especially when dealing with physically motivated predictors introduced as causal factors (Goetz et al., 2011). Moreover, Goetz et al. (2015) pointed out that the appearance of the susceptibility zonation may influence the perception of the end-user; generally, a zonation characterized by a smoothed surface is preferred and gain higher trust rather than

susceptibility maps characterized by spatial artefacts and more heterogeneous surfaces. Finally, geomorphological plausibility in terms of spatial distribution of susceptibility values is of paramount importance for the reliability of the results (Steger et al., 2016b). In the remaining part of this section, Generalize Additive Models and Random Forest will be described in detail, as these two models were employed in the present work.

Generalized Additive Models (GAM) are a semi-parametric extension of Generalized Linear Models (GLM) - or Logistic Regression in the case of a binomial response as in landslide susceptibility - introduced by Hastie and Tibshirani (1990). While GLMs can fit linear trends and are parametric in nature, GAMs extend the parametric assumption by replacing some, or all, of the parametric terms with smooth functions (Simpson, 2018). More specifically, each predictor variable in a GAM can be treated as linear (untransformed) or nonlinear (transformed by smoothing splines). Both GLMs and GAMs are used to study the relationship between a dichotomous response variable (absence/presence of landslides) and a set of explanatory variables x, both categorical and numerical (Hosmer and Lemeshow, 2000). The response variable is not modelled directly but by means of the logit of the conditional probability *p(x)*:

$$logit(x) = \ln \frac{p(x)}{1 - p(x)} = \ln(odds) \qquad \text{Eq. 2.7}$$

The odds or likelihood ratio represents the ratio between the probability p that the dependent variable is 1 and the probability 1−p that the dependent variable is 0.

In GLMs logits are modelled linearly, while in GAMs they have a more general form (Brenning, 2007, Figure 2.1):

$$logit(x) = f(x) \qquad \text{Eq. 2.8}$$

where *f(x)* is a non-linear transform of the explanatory variables and could be obtained by using smooth functions. The flexible smooths describing the behaviour of a geo-environmental predictor in a GAM is represented by a set (i.e., basis) of smaller functions (i.e., basis functions, which often are splines) that collectively contains the true smoothing function or at least a close approximation to it (Simpson, 2018, Figure 2.1).

The size of the basis is an upper limit on the expected complexity of the trend (Pya and Wood, 2016). In simple terms, each smoothing function is the sum of a certain number of basis functions, each multiplied by a coefficient, each of which is a parameter in the model. This means that a single relationship between the response and the geo-environmental predictor has several coefficients (for each

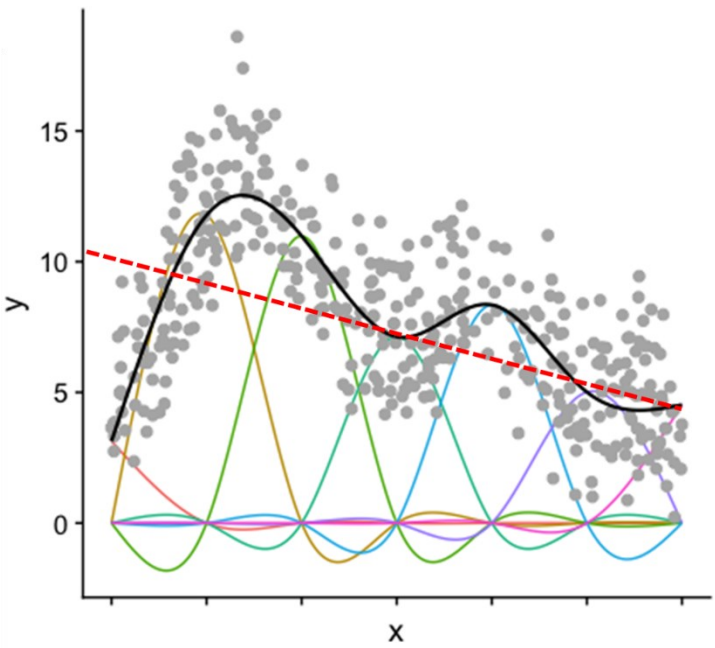


Figure 2.1 Given a set of data (grey dots) with a dependent (x) and independent (y) variable, a GLM model (red dashed line) would not capture key aspects of the data distribution, while the GAM model (black line) performed captured the non-linear behaviour of the relationship. The coloured lines represent the different basis functions constructing the smooth. (Modified from "GAM in R" by Noam Ross https://noamross.github.io/gams-in-r-course/).

basis function) plus an intercept; conversely, in the Logistic Regression framework, each independent predictor has a unique coefficient.

The goodness of fit of a GAM should be a trade-off between likelihood (i.e., the ability to reproduce the trend) and *wiggliness* (i.e., the complexity of the curve), which may lead to model overfitting. For this reason, the models are usually fit by a *penalized likelihood maximization*, in which the model likelihood is modified by the addition of a penalty for each smooth function (which penalizes its *wiggliness*). The selection of the optimal smoothing parameters could be achieved through two different approaches (Simpson, 2018): the first minimises the prediction error of the model and can be achieved by minimising Akaike's information criterion (AIC) or via cross-validation (CV) or generalised cross-validation (GCV). The second approach is to treat the smooth as a random effect, in which the smoothing parameter is treated as a variance parameter to be estimated using maximum likelihood (ML) or restricted maximum likelihood (REML Wood, 2011; Wood et al., 2016). Therefore, each smoothing function is the combination of several basis functions and smoothing coefficients and the complexity of the smoothing function in its entirety may be expressed with the effective degrees of freedom (*edf*).

In each modelling procedure, it is crucial to optimize the model through a variable selection in order to obtain a trade-off between a parsimonious model and a good performance. The selection of the important, i.e. most influencing, independent variables in GAMs modelling could be carried out through two different smoothing penalization or shrinkage approaches, able to modify smooths so they are penalized to the zero function and thereby selected out of the model (Wood, 2017): (i) the use of shrinkage smoothers, for which the smoothing penalty is modified with an additional small shrinkage

term, so that for large enough smoothing parameters the smooth becomes identically zero; (ii) the use of null space penalization, consisting in the construction of an extra penalty for each smooth, which penalizes the space of functions of zero *wiggliness* according to its existing penalties. The advantage of this approach is that it can be implemented automatically for any smooth, working as an automatic variable selection.

A common problem in modelling with GAM is linked to *concurvity*, which refers to the generalization of collinearity in the GAM setting. Concurvity ranges from 0 to 1 with 0 suggesting no collinearity, and 1 indicating that the function lies entirely in the space of one or more of the other smooth terms, meaning that one or more variables may be redundant as it can be approximated by some combination of the others and leading to unstable estimates., i.e., very sensitive to "small" variations in the model (Wood, 2017).

In the landslide susceptibility modelling field GAMs have proven to be less prone to overfitting in geomorphological modelling than Machine learning models (Hastie and Tibshirani, 1990; Brenning, 2009; Goetz et al., 2011, Petschko et al., 2012), nevertheless being still a quite flexible and interpretable model (Muenchow et al., 2012).

Random forest (RF) algorithm, firstly introduced by Breiman in 2001, is a nonparametric classification and regression supervised learning model, which constructs prediction rules without making any prior assumption on the association between the predictor features and the response variable (Probst et al., 2019). RF is considered an ensemble of decision trees, which are made of two main components: nodes and branches. Nodes could be further divided in *decision nodes,* used to make any decision and which have multiple branches, and *leaf nodes,* which are the output of those decisions and do not contain any further branches. In particular, RF is a modification of bagged decision trees, building a large collection of de-correlated trees to further improve predictive performance (Bohemke and Greenwell, 2019). The decorrelation is ensured by the random selection of a subset of predictors for nodes splitting, which also prevents over-training. Indeed, although a single decision tree is considered as a weak classifier, the combination of multiple trees in a forest is a strong classifier: every classification tree in the forest casts an unweighted vote for the sample and eventually the majority vote determines the final class of each sample of the response variable (Touw et al., 2012). In terms of landslide susceptibility, at the terminal node each tree in the forest casts a vote (e.g., landslide) and the proportion of "landslide" votes from all votes is the final predicted level of susceptibility (from 0 to 1). Among the predictor subset, the optimal predictor for node splitting (i.e., the process of dividing a node into multiple sub-nodes, according to

given conditions, to create relatively pure nodes) is usually selected by means of the decrease in Gini impurity (Boulesteix et al., 2012). This is linked to the Gini probability, i.e., the probability of correctly labelling a randomly chosen element if it was randomly labelled according to the distribution of labels previous the new split. The more often a predictor is selected as the best splitter in the random subset, the higher would be its "variable importance".

Another important feature of RF is that they are based on random bootstrap samples selected with replacement (Efron and Tibshirani 1986): differently from cross-validation (e.g., k-fold, holdout or leave-one-out), bootstrapping allows a data point to be selected several times in a particular training subset, nevertheless maintaining approximately the same distribution of values as the original data set. The original observations not contained in a particular bootstrap sample are the so called out-of-bag (OOB), which is used to validate the model (Bohemke and Greenwell, 2019). A summary of how RF works is presented synthetically in the workflow in Figure 2.2.

Lagomarsino et al. (2017) well summarized the advantages of the RF technique: i) it can handle both categorical and continuous variables; ii) prior statistical assumption on data are not required; iii) it can account for predictor variables mutual interactions and nonlinearities; iv) it is robust with respect to changes in the dataset and to noise features; v) validation on OOB data minimizes overfitting; vi) it can handle a large number of predictor variables.

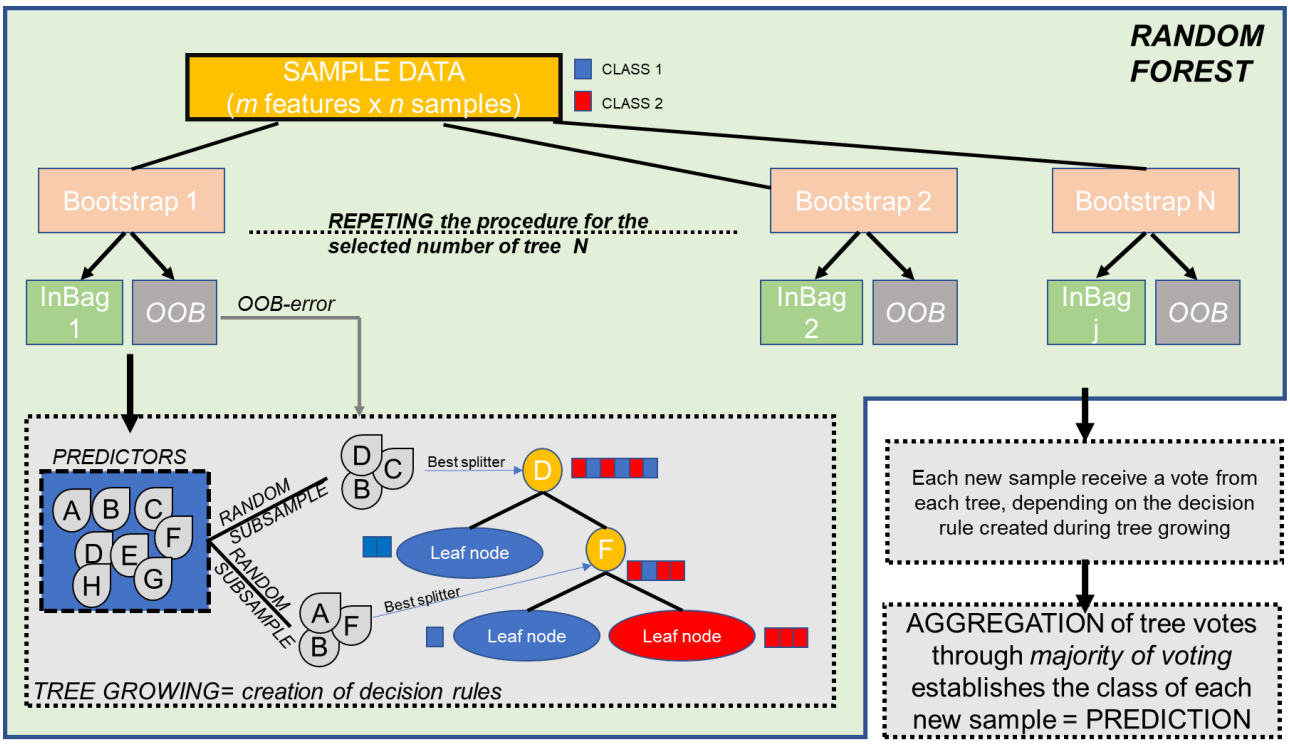


Figure 2.2 Workflow and steps of the Random Forest algorithm.

An important step in RF is the hyperparameter tuning phase, i.e., finding an optimal combination of model parameters that should be a trade-off between performance and runtime (Bohemke and Greenwell, 2019). However, the risk in selecting a parameter set leading to complex rules could end up in model overfitting, too specific to the training data but worse performing on the validation set; however, this also depends on the objective of the model, i.e., if it made for future prediction or for explaining the relevance and the behaviour of the candidate predictor variables (Probst et al., 2019; Shmueli, 2010). Several authors investigated the effects of hyperparameter tuning on model behaviour, performance and runtime (Segal, 2004; Lin and Jeon, 2006; Strobl et al., 2007; Grömping, 2009; Martínez-Muñoz and Suárez, 2010; Goldstein et al., 2011; Janitza et al., 2016; Wright and Ziegler, 2017).

According to Probst et al. (2019), who give a comprehensive review of hyperparameters and tuning strategies, the most important hyperparameter are:

- mtry – It controls the number of candidate variables randomly selected at each split during trees growing. Even if the default value for classification set to mtry=√p usually performs well, this parameter selection should be a trade-off between stability and accuracy of each single tree. Low values of mtry lead to less correlated trees, favouring stability during bagging; however, it could lead to trees performing on average worse, because built on potential suboptimal or non-important variables. When there are few relevant predictors a higher value of mtry tends to perform better because the chance to select the strongest signal is higher. However, the less influential variables (but still important for some sub-groups of samples) could be masked by the strongest ones.
- Sample size –This parameter determines how many observations are drawn for the training of each tree. Theoretically, bootstrapping provides for the100% of the observations sampled with replacement. However, decreasing the sample size leads to more diverse and thus more de-correlated trees (i.e. higher difference in OOB samples in each tree), which nevertheless could cause a decrease in the accuracy of the single trees since fewer observations are used for training.
- Node size – It controls tree complexity specifying the minimum number of observations in a terminal node (i.e. increasing node size means a decreasing tree depth and complexity). In addition, along with mtry, it is the parameter with the highest control on runtime.

- ntree – The number of trees in the forest, although not technically a hyperparameter, controls error rate (i.e. OOB error) and variable importance estimate. Even if with a consistent impact on runtime, more trees are always better (Probst and Boulesteix, 2017).

The selection of the optimal hyperparameters could be performed both by k-fold cross validation (Seibold et al., 2018) and by the OOB error minimization (Probst et al., 2019). Hyper-tuning strategies are different, from the simple grid or random search, in which all possible combinations are respectively tested from a given discrete parameter space or are randomly selected from an hyperparameter space to more sophisticated techniques as sequential model-based optimization (Hutter et al., 2011), which relies the future hyperparameter set to be tested on the results of the previously evaluated ones.

Random Forest technique has been widely used in many applications, including landslide susceptibility (Brenning 2005; Catani et al., 2013; Paudel and Oguchi 2014; Segoni et al., 2015, 2020; Youssef et al., 2016; Lagomarsino et al., 2017). Also, it often showed better quantitative performances when compared with other more traditional methodologies such as logistic regression (e,g,, Trigila et al. 2013; Goetz et al., 2015), as machine learning algorithms are specifically developed to predict with high accuracy complex interactions (Elith et al., 2006). However, RF, as the majority of machine learning algorithms, is considered as a black-box method, thus the interpretation of the model behaviour is more complex (Elith and Leathwick, 2009). More traditional methods, such as logistic regression and Generalized Additive models, provide more easily interpretable results, especially in presence of physically motivated predictors, representing processes directly associated to landslides (Goetz et al., 2011, 2015; Camera et al., 2021). Moreover, machine learning derived susceptibility maps could result more likely affected by spatial artefacts and pixelation, affecting in a negative way the clear definition of the most hazardous zone and the potential user perception of the method (Brenning, 2005, Brenning, 2012b, Goetz et al., 2015).

In this framework, a possible way to overcome the complexity of RF, making this technique more suitable for interpretation and for the understanding of predictors behaviour, consists in the use of the SHapely Additive exPlanation (SHAP) framework introduced by Lundberg and Lee (2017). The SHAP are based on Shapley values, initially proposed by Shapley (1953), a game theory concept: if the prediction is considered as the game "payout", and the feature (i.e., model predictor) values of the data sample is considered as a "player", Shapley values explain how to fairly distribute the "payout" among the *N* features. In other words, the Shapley value is the expected average marginal contribution of adding (or removing) a player *i* to the game; "average" is referred to the fact that the marginal value is averaged considering the contribution of the addition (or exclusion) of the player *i* to all the possible combinations

(or subsets) of the N-*i* remaining players in the game. A Shapley value could increase (positive effect) or decrease (negative effect) the prediction of each data sample, contributing to the deviation from a baseline (i.e., the average of all predictions or the average model output value over the training dataset). These forces balance each other, resulting in the final prediction of an instance (i.e., a data sample) (Du M. et al., 2019; Molnar, 2019).

Shapely values are exploited in the SHAP framework to calculate feature importance and are computed by sequentially introducing each feature (i.e., predictor) into a conditional expectation function $fx(S)$ of the model's output, attributing the change in expectation to the feature introduced, then averaging the process over all possible feature orderings so that the features are fairly compared. This is crucial to maintain consistency, since the order in which a model sees features can affect its predictions (Lundberg et al.,2020). Two main features make SHAP values particularly suitable for RF models interpretation (and in general machine learning models interpretation): (i) global interpretations are consistent with the local explanations, since the Shapley values are the "atomic unit" of the global interpretations. In other words, SHAP values allow both analysing each predictor variable effect, magnitude and direction on model outputs at a global level and the specific effect of each predictor on each sample at the local level (e.g., landslide and non-landslide points). (ii) they satisfy the additive property of game theory, which guarantees that the average of the Shapley values of each feature (predictor) extracted from the individual trees corresponds to the Shapely value of the feature in the forest (Molnar, 2019).

SHAP values implementation with tree-based machine learning techniques (e.g., Random Forest and XGBoost) have recently gained popularity in different fields, such as medicine (Lundberg et al., 2020; Li et al., 2020), finance (Mokhtari et al., 2019), engineering (Mangalathu et al., 2020) and even in the field of geosciences (Lubo-Robles et al., 2020; Voltolina, 2021). However, until now, the usage of the SHAP framework is very rare in the landslide susceptibility field, with a unique recent work by Can et al. (2021) found in the literature.

2.3.6. Validation techniques and uncertainties evaluation

Susceptibility model performance is evaluated by means of different indices and metrics (Guzzetti et al., 2006; Frattini et al., 2010; Rossi et al.,2010). The entire process of performance evaluation relies on the preliminary individuation of a training set (i.e., a set of landslides employed to construct the susceptibility model) and a test set (i.e., a set of landslides used to verify independently the model) (Guzzetti et al., 2006). This leads to the necessity of a preliminary distinction between model fitting (i.e., success rate) and model prediction (i.e., prediction rate) performance (Reichenbach et al., 2018). The

former refers to the ability of the model to match the distribution of landslides used to train the model and to calibrate the functional relationships with the geo-environmental predictors; the latter refers to the ability of the model to predict a set of landslides that are independent and were not used to construct the model (i.e., a test set). It is important to highlight that any statistical classification provides better results on the training set and performs less efficiently when applied to the validation set (Michie et al., 1994; Guzzetti et al., 2006).

The identification of these two distinct sets is typically carried out through random, spatial and temporal selection strategies. Some researchers focused on the influence of different training-test sets selections on model performance both through random (e.g., Sameen et al., 2020; Kalantar et al., 2018) and spatial guided selection (e.g., Erener et al., 2020) strategies. The temporal selection strategy specifically requires either an event-based (e.g., Knevels et al., 2020) or a multitemporal inventory (e.g., Samia et al., 2020; Lombardo et al., 2020). The most used ratios to split the available sample into training and test subsets are 70/30 and 80/20 (Reichenbach et al., 2018; Nguyen et al., 2021). A 90/10 ratio may also be used and proved to be reasonable especially in very complex modelling frameworks (e.g., Lombardo et al., 2020). Frequently, single hold-out methods are used to split the data in a training and test sample. This procedure however leads to a single estimate of the model performance, without providing a measure of the metric precision; the performance indeed depends on the sample (randomly, spatially or temporally selected) and on the peculiar characteristics of the test set (Petschko et al., 2014; Sameen et al., 2020; Kalantar et al., 2018; Erener et al., 2020). This issue could be addressed by using repeated k-fold cross-validations (Petschko et al., 2014; Goetz et al., 2015; Steger et al., 2016a,b; 2017; 2021; Camera et al., 2021, Lombardo et al., 2020). In this framework, different subsets (i.e., folds) are used either as the train and the test set, thus using the entire dataset both for evaluating the fitting and the prediction performance of the model and reducing sampling variability (Brenning 2012a,b; Petschko et al., 2014). Thus, instead of having just one performance metric for the training and for the test set, a range of independent performance estimators would be obtained (in a quantity equal to the number of selected folds times the number of repetitions). Moreover, this procedure may be applied both with a random subsampling of folds and through a spatial subsampling (i.e., spatial cross-validation based on k-means clustering of point coordinates) (Ruß and Brenning, 2010). A leave-one-out cross validation has shown to be more appropriate when dealing with the temporal selections of folds, in this case representing particular time intervals (Lombardo et al., 2020).

The most frequently adopted metrics to evaluate model performance are (Chung and Fabbri, 2003; Remondo et al., 2003; Melchiorre et al., 2011; Frattini et al., 2010; Corominas et al., 2014; Reichenbach et al., 2018, Triglia et al., 2015):

- The Success and Prediction rate curves, calculated on the training and test subsets, respectively. They display in descending order the susceptibility level on the x-axis and the cumulative percent of landslide occurrence on the y-axis.
- Contingency tables or "Confusion matrix" (Jollifee and Stephenson, 2003), displaying the amount of True Positives (TP), i.e. the match between observed landslides and predicted unstable terrain units, True Negatives (TN) i.e., the match between observed non-landslides and predicted stable terrain units, False Positives (FP), i.e., the mismatch between not-observed landslides and predicted stable terrain units and False Negatives (FN), i.e., the mismatch between observed landslides and predicted stable terrain units. Triglia et al. (2015) argued that different mismatches may cause different issues: FP may lead to a loss of economic value of some areas, while FN may determine a not socially acceptable cost in case of casualties and damage to exposed elements.
- the ROC curve (Receiver Operating Characteristics Curve, Green and Swets, 1966;) plot, for different susceptibility threshold values (between 0 and 1, and usually set to 0.5) the TP rate (i.e., TP/TP+FN and also called Sensitivity) and the FP rate (i.e., FP/FP+TN and also indicated as 1-Specificity).
- The Area Under the ROC Curve (AUC or AUROC), which varies from 0.5, indicating a pure chance agreement between predictions and observations to 1, indicating a perfect model capability in discriminating landslides.

To recall the previous Sections, the performance metrics are frequently used to evaluate different classification algorithms (e.g., Goetz et al. 2015) and ensembles (e.g., Di Napoli et al., 2020), to optimize predictors selection (Conoscenti et al. 2016; Amato et al., 2020; Camera et al., 2021), to evaluate the spatial transferability of modelling results (Petschko et al. 2014; Lombardo et al. 2014), to quantify the influence of sample sizes (Petschko et al. 2014; Hussin et al. 2016) and the effect of sampling strategies (Petschko et al.,2013, 2016; Conoscenti et al. 2016; Hussin et al. 2016; Otzurk et al., 2020) and to investigate discrepancies between validation results and geomorphological plausibility of the susceptibility maps (Steger et al., 2016b). Regarding the plausibility of the susceptibility zonation, it is of vital importance; even statistically high performing models may suffer from geomorphological errors, which, in a risk management perspective, are often more severe than a model with a lower quantitative

performance but a more physically reliable outputs (Carrara et al., 1991; Reichenbach et al., 2018). Also, it should be remembered that models with very similar performance levels may produce very different susceptibility maps (Triglia et al., 2013; Sterlacchini et al., 2011; Goetz et al., 2015).

Uncertainties intrinsically characterize every modelling effort to reproduce a natural system, a natural process or a natural property, including landslide susceptibility. Uncertainties may be distinguished in epistemic and aleatory (Oberkampf et al., 2004; Roy and Oberkampf, 2011). Epistemic uncertainty may be referred as the imperfect understanding of the studied phenomenon and could be reduced by improving the knowledge of the process and the governing parameters (Ardizzone et al., 2002; Petschko et al., 2014). Aleatory uncertainty is unavoidable as it is linked with the inherent variability of the natural systems and its randomness (Rougier, 2013). The evaluation and implementation of the epistemic uncertainties in landslide modelling and their quantification gained attention only in recent years (Reichenbach et al., 2018), but a standardize procedure is still missing. A benchmark work on this topic was carried out by Petschko et al. (2014). Authors discussed the different sources of epistemic uncertainty in landslide susceptibility modelling and recognized that uncertainty may arise during different phases of the modelling process: uncertainty related to the input data, uncertainty related to the model form and uncertainty on model predictions.

The first uncertainty source is related to input data, both from the inventory and from the geo-environmental predictors. The estimation of landslide inventory completeness and the effects of positional accuracy is a topic addressed by several authors (Zezere et al., 2009; Petschko et al., 2013, 2016; Steger et al., 2017; Conoscenti et al. 2016; Hussin et al. 2016; Otzurk et al., 2020). Geo-environmental predictors related uncertainty is conversely analysed in terms of the effects of DEM resolution (e.g., Chen Z. et al., 2020; Rabby et al., 2020a) or of elevation synthetically generated errors (e.g., Murillo and Hunter, 1997; Qin et al., 2013). More recently, Huang et al. (2021) investigated the influence of selecting the attribute interval numbers (AINs) when reclassifying continuous predictors in categorical variables (a step sometimes required by some modelling algorithms).

The second source of uncertainties lean on the model itself and may be addressed by different indices and validation procedures. For instance, preferring a cross-validation instead of an holdout validation allows to obtain the precision of the performance (Brenning 2012a,b); opting for a spatial cross-validation provide for spatial transferability measures of the model on surrounding areas, leading to physically sounder results (Petschko et al., 2014); implementing in the cross-validation phase the

calculation of indices such as variable selection frequency (Goetz et al., 2011) and thematic consistency (Petschko et al., 2014) allows to verify the robustness of the modelled functional relationships.

Finally, prediction uncertainty derives from the fact that the individual probability values obtained as output are, in reality, estimated conditional mean values of the predicted probability (Hosmer and Lemeshow, 2000). Indeed, each susceptibility value carries with it a prediction uncertainty (i.e., a probability range), which may be more or less wide. The analysis of this standard error was addressed by several researchers, essentially working with Logistic Regression and GAMs (e.g., Guzzetti et al., 2006; Rossi et al., 2010; Petschko et al., 2014). However, with the establishment of more complex modelling algorithms belonging to the Machine Learning category, this statistical output (i.e., the probability range) may be more difficult to obtain. To encompass this, Ensemble Modelling is becoming popularly employed in classification problems and exploited to reduce uncertainty by combining predictions yielded by multiple algorithms (e.g., Umar et al. 2014; Youssef et al. 2015; Pham et al. 2017; Kim et al. 2018; Bueechi et al. 2019). More specifically, uncertainty is often assessed through the calculation of the Coefficient of Variation (CV), which is the standard deviation of probabilities obtained for each pixel, due to the agreement (or disagreement) level in predicting susceptibility amongst the models used to produce the Ensemble (Kim et al, 2018; Di Napoli et al.,2020).

Chapter 3

Valchiavenna Case Study

3.1. Research question and specific objectives

In geologically complex, large alpine areas, such as Valchiavenna (SO, Central Italian Alps), the challenge is to identify differences in failure probability amongst adjacent cliffs, and to understand why different portions of the same rock mass behave differently (Matasci et al., 2018). Where the climatic and topographic contexts are similar, variability in rockfall activity should be investigated in terms of geomechanical characteristics, stress state, and variations in hydrogeologic conditions of the rock walls (Coe and Harp, 2007).

Regional-scale susceptibility approaches that only include topographic predictors may suffer from a too strong generalization of the processes, as these phenomena reflect a complex interplay of numerous processes acting at different spatial-temporal scales (Messenzehl et al., 2017). Nonetheless, including other types of predictors, relevant to the landslide type and volume, is challenging. Geo-structural and geomechanical properties are difficult to interpolate (Reichenbach et al., 2018), primarily due to the rare availability and suitable coverage of these data over large areas. When available, these types of datasets are not suitable for interpolation purposes, because they are prepared in relation to local problems or clustered close to roads and infrastructures, thus not representative of the properties' variability. Therefore, strategies for the design of optimal sampling schemes (i.e., efficient and effective in terms of time, budget, and quantity/quality of sampled points) to update or even create geological-environmental datasets become necessary. Several of such options are available in the digital soil mapping literature (see Brus, 2019 for a comprehensive review) and some of them are presented in Section 2.1.

Rock mass fracturing degree, weathering and hydrogeological properties are considered fundamental parameters describing rock-mass mechanical conditions and quality, and therefore included in the most widely used Rock Mass Classification systems (e.g., Rock mass Rating-RMR by Bieniawski, 1973; Q-System by Barton et al., 1974). The presence of discontinuities of different origin in rock masses creates an anisotropic strength dramatically affecting rock slope stability (Coe and Harp, 2007), and deformability is more strongly influenced by joints than by intact-rock properties (Jaboyedoff et al., 2004). Bedrock in-situ stresses induced by the complex interplay between exhumation, ongoing tectonic processes, glacial unloading and micro-crack nucleation, play a key role in shaping the morphology of Alpine valleys, driving the fracturing pattern of rock masses (Leith, 2012). Owing to these complex

interacting mechanisms, the role of fracturing grade, at least when working at the rockfall scale, should not be reduced to the traditionally used distance-from-fault predictor, being actually the spatial variation of the geometrical characteristics of jointing the key issue for rockfall distribution at the regional scale (Wang et al., 2021).

Weathering is another factor affecting the instability of rock slopes and rock mass quality (Eberhardt et al., 2005; Ceryan et al., 2008; Zimmer et al., 2012; Miščević and Vlastelica, 2014; Krautblatter and Moore, 2014). The degree of weathering is linked to slope strength and acts as a quasi-static preparatory factor for instability, whereas weathering rate relates to stresses acting on the slope. More complex, non-linear, interdependent relationships between weathering and slope instability could also originate within the rock mass system (Viles et al., 2013; Krautblatter and Moore, 2014).

In crystalline rocks, where intact rock permeability is negligible, the fracture system dominates fluid flow and the permeability of the rock mass system (Baghbanan and Jing, 2007). Rock mass permeability influences slope stability in a complex dual relationship: time-dependent crack damage and fracture growth enhance rock mass permeability (Riva et al., 2017; Grämiger et al., 2017), which in turn controls infiltration, discharge, and pore pressure changes. Due to the difficulties in quantifying rock mass hydrogeological conditions objectively, they are usually included in a qualitative way in most of the rock mass classification systems, especially when applied to slope stability problems (Pantelidis, 2009).

Focusing on the acquisition of good quality and relevant geo-environmental predictors for landslide susceptibility modelling, rather than experimenting with new modelling techniques, is a challenge that needs to be addressed (Reichenbach et al., 2018). Embracing this suggestion, the overarching goal underlying the Valchiavenna Case Study was to test the influence of some geomechanical properties on rockfall susceptibility in an Alpine environment. Specific objectives, summarized in Figure 3.1, were:

(i) the update and review of an available geomechanical dataset for Valchiavenna through the optimal selection of additional sampling points and the execution of the geomechanical surveys;

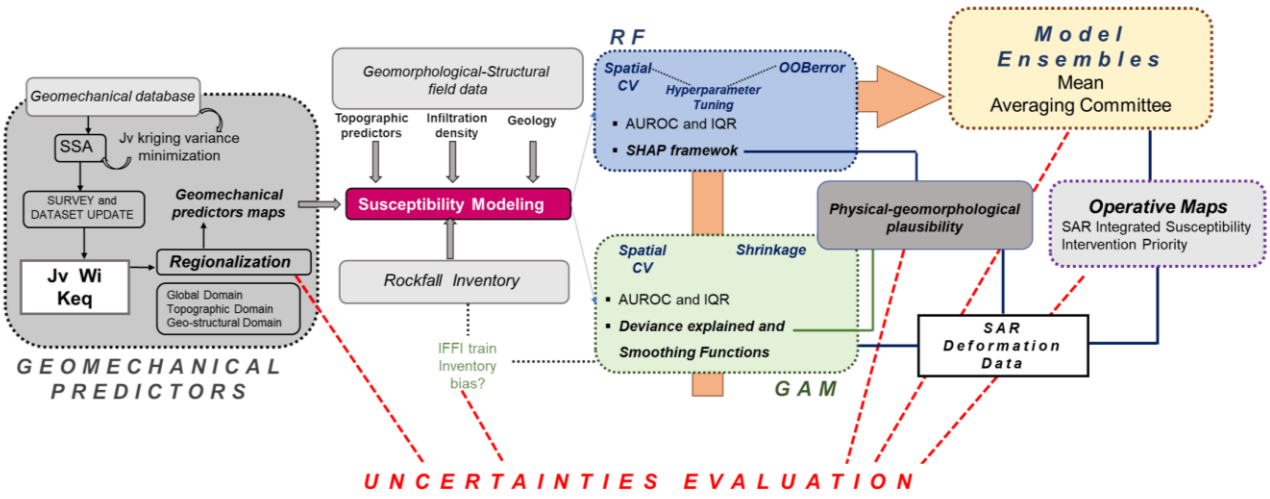
(ii) the calculation of three geomechanical properties, derived from the post-processing of geomechanical surveys - namely Joint Volumetric Count (Jv), rock-mass weathering index (Wi) and rock-mass equivalent permeability (Keq), - and their regionalization over the study area, comparing different regionalization techniques and domains;

(iii) the implementation of the obtained regionalized geomechanical properties as predictors in a rockfall susceptibility model, performed both using Generalized Additive Models (GAM) and Random Forest (RF);

(iv) the integration of the GAM and RF resulting maps, with the aim of producing a quantifiable and spatially distributed uncertainty measure of rockfall susceptibility;

(v) the integration of the rockfall susceptibility maps with satellite-derived ground deformations, to produce easily readable operational tools, useful for environmental and risk management planning.

Moreover, two threads guided the entire procedure: (i) the recognition of the uncertainties arising from each step of the susceptibility model building, (ii) the evaluation of the geomorphogical-geological plausibility of the outputs. Uncertainties may arise from the rockfall inventory (Zezere et al., 2009; Petschko et al., 2013, 2016; Steger et al., 2017; Conoscenti et al. 2016; Hussin et al. 2016; Otzurk et al., 2020). Uncertainties are also linked to the inherent flexibility of the modelling algorithm in explaining predictors' behaviour (Umar et al. 2014; Youssef et al. 2015; Pham et al. 2017; Kim et al. 2018; Bueechi et al. 2019). Moreover, uncertainties quantification may be exploited to provide products necessary for environmental planning and risk management (Petschko et al., 2014; Ciampalini et al., 2016; Di Napoli et al., 2020). Plausibility of the outputs should be taken into account during the whole process, from the geomechanical predictors regionalization to their behaviour in the susceptibility model, as well as the



ACRONYMS

SSA Spatial Simulated Annealing
Jv Rock mass Joint Volumetric Count
Wi Rock mass Weathering Index
Keq Rock mass Equivalent Permeability
IFFI Italian National Landslide Inventory
GAM Generalized Additive Model
RF Random Forest
AUROC Area Under the Receiver Operating Characteristic Curve
Spatial CV Spatial Cross Validation
OOBerror Out of Bag error
SAR Synthetic Aperture Radar

Figure 3.1 Workflow representing the main phases of the project.

interpretability of the output maps (Brenning, 2005, 2012, Goetz et al., 2011, 2015; Steger et al., 2016b, 2021; Camera et al.,2021).

3.2. Study area

Valchiavenna (northern Italy, Province of Sondrio) is a structural-glacial valley of about 275 km², located in the Central Alps (Figure 3.2a). It is characterized by the convergence of two orthogonal tributary valleys in correspondence of the town of Chiavenna: San Giacomo Valley (N-S), and Bregaglia Valley (E-W), both connecting Italy to Switzerland. These two valleys differ in terms of geo-structural setting and slope morphology (Figure 3.2b and Figure 3.2c).

The San Giacomo Valley is predominantly characterized by the sub-horizontal contact between the tabular gneissic bodies of the Tambò and Suretta Pennidic nappes. Both nappes are characterized by a

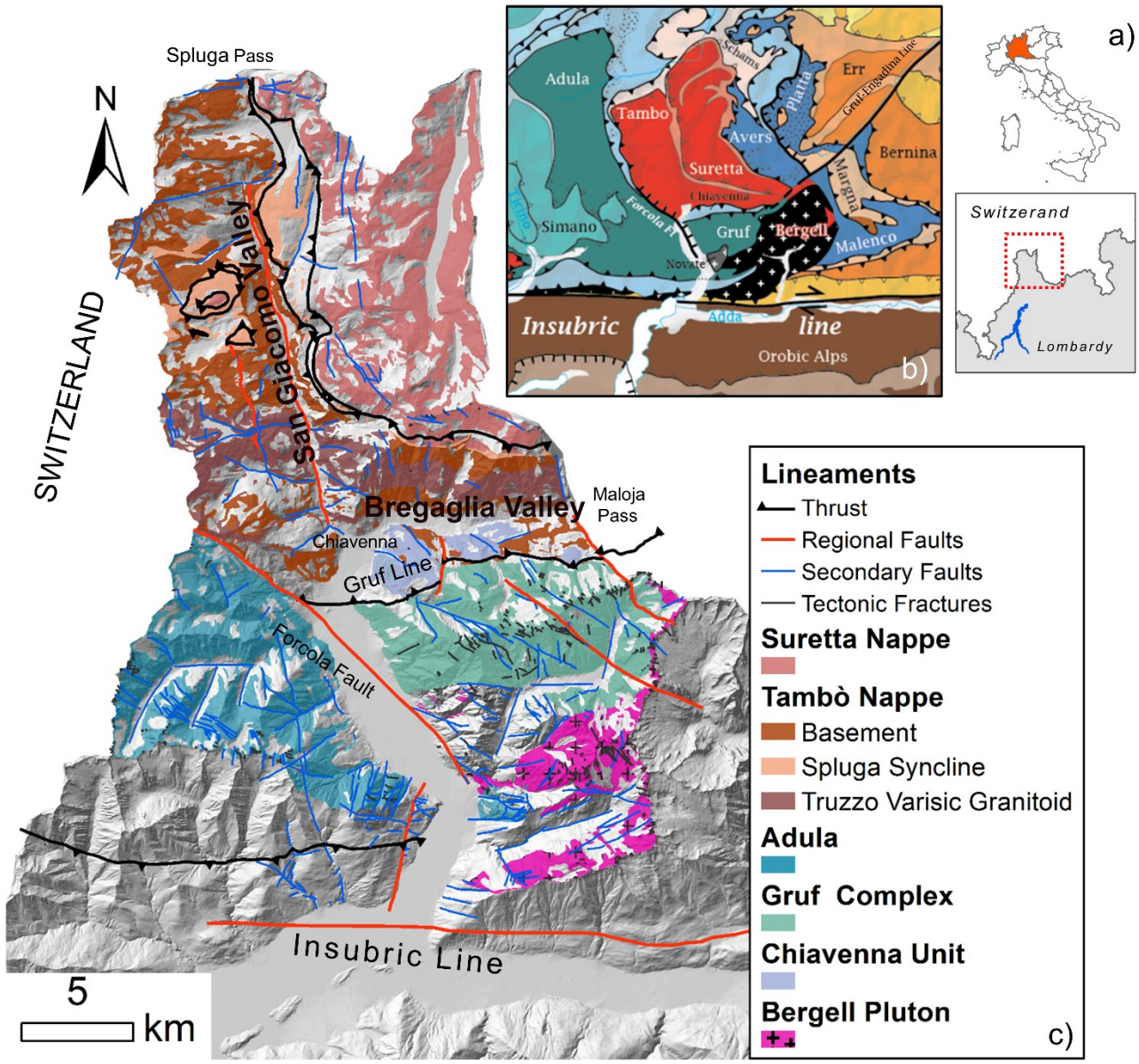


Figure 3.2 . a) Location of the study area (Lombardy region, Italy); b) geo-structural framework of the study area in the alpine chain (extracted from Tectonic framework of the Alps, Bousquet et al., 2012); c) Detailed geological and structural map of the study area (modified from Montrasio &Sciesa, 1988).

gently NE-dipping shistosity and by similar polycyclic and poly-metamorphic paragneiss lithologies, intercalated with orthogneiss and ultramafic layers (Montrasio & Sciesa, 1988; Nussbaum et al., 1998). The basement of Suretta Nappe could be further subdivided into two main units: the heterogeneous polymetamorphic rock assemblage called Stella-Timun Complex, outcropping in the Southern part of the Valley and the Permian igneous body of the Porphyric Roffna Complex (Scheiber et al., 2012). The crystalline basements of the two nappes are separated by the Spluga Syncline Permo-Mesozoic metasedimentary cover, showing the typical facies of the Brianconnais domain (Baudin et al., 1995) with metapelites, carbonates and quartzites. These lithologies outcrop more frequently in the northern part of the San Giacomo Valley, resulting in a wide, open morphology of the valley, with large flat surfaces at mid-high elevations (1500-2000 m a.s.l.). Conversely, the southern part of the San Giacomo Valley is shaped by the presence of the Varisican Truzzo Metagranite, intruded into the Tambò Nappe, resulting in a narrow valley with steep slopes (Tantardini et al., 2013). The Alpine metamorphic grade increases from the top of the Suretta nappe to the bottom of the Tambò nappe and from the North to the South of nappes from greenshist facies to amphibolite facies (Baudin & Marquer, 1993). The structural contact between the Tambò and Suretta nappes extends through the northern slope of the Bregaglia Valley, where it is clearly visible at an elevation of around 2000 m a.s.l. The southern slope, in contrast, is characterized by the presence of the Penninic granulite-migmatite Gruf Complex in structural contact with the ultra-mafic Chiavenna Unit, marked by the sub-vertical mylonitic zone called Gruf Line (Galli et al., 2013). In its south-eastern sector, the Gruf Complex is intruded in a compressional tectonic regime (Berger et al., 1996) by the Periadriatic Bergell Pluton (Tibaldi and Pasquarè, 2013). The Chiavenna Unit is interpreted as the result of the youngest basic oceanic magmatism in the Alps (Liati et al., 2003) originating from stripped subcontinental mantle tectonically exposed during progressive oceanization (Huber and Marquer, 1998). This Unit consists mainly of metaperidotites, amphibolites, metagabbros, and rare carbonate rocks and is tectonically located above the Gruf Complex and below the Tambò Nappe (Liati et al., 2003).

The structural framework of the entire area is influenced by some main regional tectonic alignments: a WNW-ESE system related to the Insurbric Line, a NW-SE system linked to the Forcola Fault, and a NE-SW system associated with the Engadine Line (Ferrari et al., 2014). On the southern slope of the Bregaglia Valley, the Gruf Line, interpreted as the brittle-ductile elongation of the Engadine Line (Wenk, 1984), dominates the structural setting. Moreover, a bundle of tensional joint sets, parallel to the valley axes, are locally observable as a result of deglaciation related stress release (Ferrari et al., 2014), superimposed on pre-existing weakness and acting in association with local stress state. In San Giacomo Valley, the structural setting of the two Pennidic nappes is the result of five recognizable superimposed

deformation phases (Wiederkehr et al., 2008; Scheiber, 2012). More specifically the main regional shistosity marking the contact between the Suretta and Tambò Nappes is ascribable to the second deformation phase, which induced the most penetrative ductile structures. These structures were cut by subsequent ductile detachment zones and by late and post-alpine brittle-ductile deformation processes. More recently, two late deformation phases overprinted and steepened the previous structures, producing an extensive fracturing pattern expressed by normal faults (Ferrari et al., 2014). The structural setting of Bregaglia Valley is dominated by the three main ductile shear zones of the Gruf Complex, carrying identical mineral association suggesting a contemporaneous development (Galli et al., 2013): (i) a first group striking ENE-WSW and steeply dipping towards NNW, also marking both the main foliation and the structural contact between the Gruf Complex and the Chiavenna Unit; (ii) a second group steeply North dipping with a dextral normal sense of shear; (iii) a third group of NW-SE striking shear zones steeply dipping towards NE with a dextral normal sense of shear.

The interplay between the above-mentioned complex and polycyclic tectonic processes and glacial cyclical de-buttressing (Grämiger et al., 2017), following the retreat of the Engadine-Bregaglia and San Giacomo glaciers (Tantardini et al., 2013), is crucial in defining the present-day slope dynamics in the study area (Figure 3.3). Both the San Giacomo and the Bregaglia Valley are characterized by several DSGSD (Deep Seated Gravitational Slope Deformations) with different "maturity" and morpho-structural evidence (Tantardini, 2016). Associated to these wider and slower processes, secondary more rapid phenomena such as rock avalanches, rockslides and rockfalls could occur. In the study area, the most ancient known events of such type are the Cimaganda rock avalanche in the San Giacomo Valley, which probably occurred in the IX century (Mazzoccola, 1993), and the 1618 Piuro landslide on the southern slope of Bregaglia Valley (Schuster and Highland, 2007). These big events aside, the community living in Valchiavenna frequently faces several smaller instabilities and their associated damages and risks. The most recent events are the Cimaganda reactivation in 2012 (Morcioni et al., 2020), the Cengalo rockslide-rockfall in 2017 (Mergili et al., 2020) and the Gallivaggio rockfall in 2018 (Carlà et al., 2019).

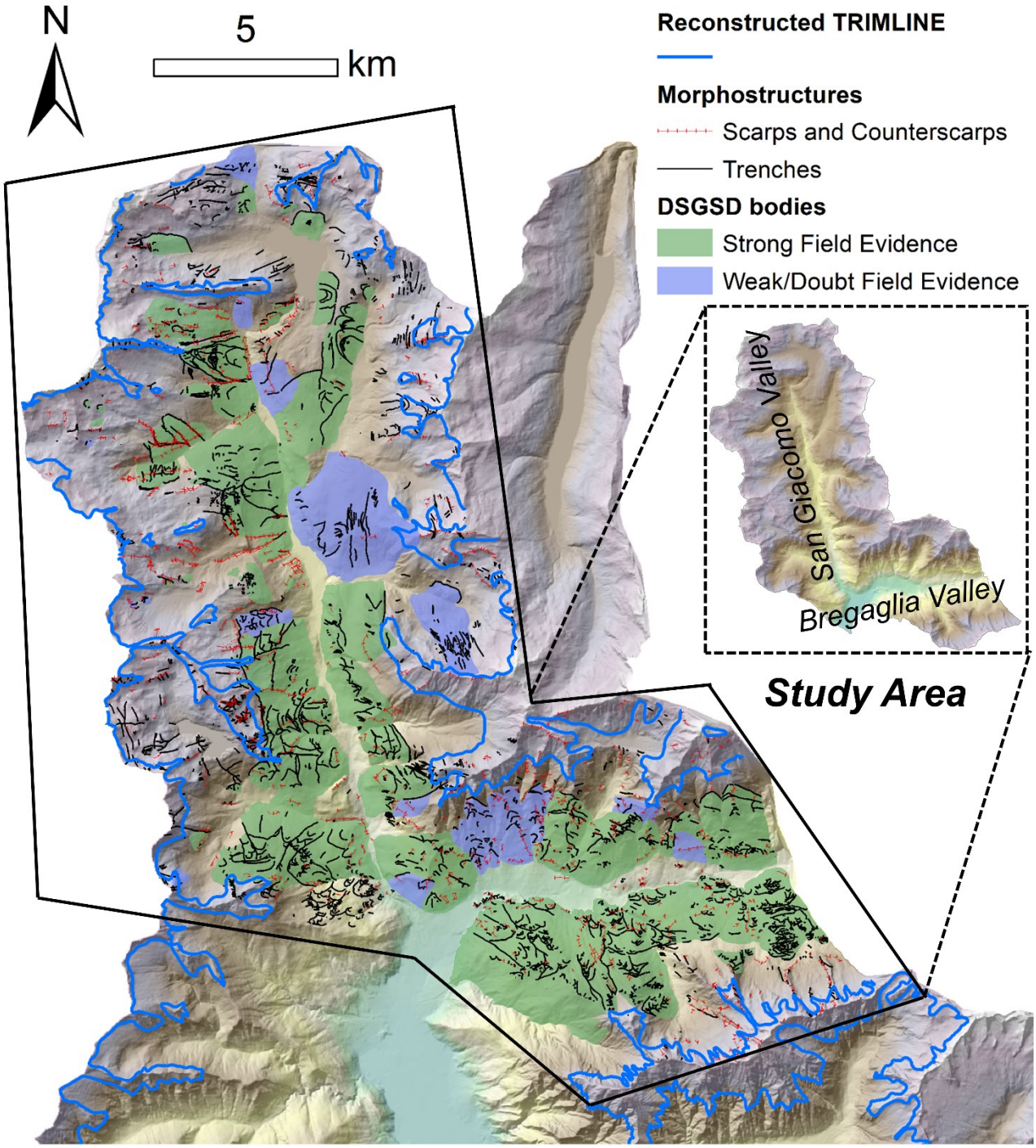


Figure 3.3 Geomorphological framework of the study area: principal geomorpho-structures, DSGSDs and reconstructed trimline, all based on field-survey evidence (data from the Valchiavenna Project, Tantardini et al., 2013 ad Tantardini, 2016).

3.3. Data

3.3.1. Land surface data

In this study the Digital Terrain Model (DTM) made available by Regione Lombardia (http://www.geoportale.regione.lombardia.it/) was used. It has a 5 m x 5 m horizontal resolution and is the result of several airborne LiDAR survey campaigns (2008-2009; 2010-2011; 2013-2015).

The geological map for the study area was available at a 1:50,000 scale (Crop Project, Montrasio and Sciesa, 1988), and reporting the main geo-structural domains together with a more detailed description based on lithologic and mineralogic characteristics.

SqueeSAR™ data (TRE Altamira) were available for the study area and consist of permanent (PS) and distributed (DS) scatterers, reporting line-of-sight (LOS) displacement rate. Data are available in both ascending (track=15, mean LOS=40.64°) and descending (track=66, mean LOS=36.94° and track=168, mean LOS=42.4°) geometries. In detail, for this study data derived from Sentinel 1A/B radar from October 2014 to February 2021 were made available by Regione Lombardia within the framework of the ongoing AMALPI Interreg Project (https://progetti.interreg-italiasvizzera.eu/it/b/78/alpiinmovimentomovimentonellealpipiuro).

3.3.2. Rockfall data

Rockfall data were the result of the integration of two different rockfall inventories: (i) the freely available IFFI dataset (Inventario Fenomeni Franosi in Italia, https://www.progettoiffi.isprambiente.it/) where rockfalls are partly collected as polygons reporting both the source and the deposits area and as points labelled as "historical events". For the polygon type, the rockfall source points were extracted from the publicly available dataset of the project ROCKtheALPS 2019 (alpine-space.eu/projects/rockthealps/); (ii) a geomorphological dataset (from here on called UNIMI inventory), comprising several additional rockfall scarps and deposits and covering also remote areas. The latter was extracted from a detailed geomorphological-structural map of the area, derived from field surveys at a 1:10,000 scale, integrated with remote sensing, carried out from 2002 to 2020 in the framework of the Valchiavenna Project (Sfondrini and Pasquarè, 2011), several MSc theses, the Ph.D. studies of Tantardini (2016), and within the ongoing A.M.AL.PI.18 Interreg Project.

3.3.3. Geomechanical data

A dataset of 128 geomechanical surveys was available for the study area (database of the geoengineering research group of the Dept. of Earth Sciences, Università degli Studi di Milano). Each record was

acquired according to the International Society of Rock Mechanics (ISRM) Suggested Methods (ISRM, 1978), including both primary variables (e.g., Joint Roughness Coefficient JRC, Schmidt Hammer rebounds, joint spacing, orientation, aperture, persistence) and derived rock mass quality indices (e.g., RMR, GSI). This dataset was the result of several detailed geomechanical campaigns carried out by different surveyors (mainly MSc and PhD students) since 2000 and with different underlying objectives. For this reason, although quite rich, the dataset came with different levels of information and completeness and is clustered along the main roads (i.e., accessible areas).

3.4. Methods

3.4.1. Geomechanical properties relevance for rockfalls and their calculation

Rock mass systems are defined as DIANE i.e., discontinuous, inhomogeneous, anisotropic and nonlinearly elastic materials (Hudson and Harrison,1997; Krautblatter and Moore 2014), recording a complex suite of thermal, hydrogeological, mechanical and chemical processes, which have been interacting for millions of years (Jaboyedoff et al., 2011; Krautblatter and moore 2014). The great variability in the spatial distribution of rock slope instabilities reflects rock mass local conditions. These can vary in terms of strength and deformability of the intact rock, as well as in terms of fracture network physical-mechanical properties (Loye et al., 2012; McColl, 2012). Indeed, rock mass conditions and resulting mechanical behaviour derive from the interaction of topographic, tectonic, glacial loading-unloading and exhumation generated stresses, acting at several spatial and temporal scales and preferential directions (Ballantyne, 2002; Jaboyedoff et al., 2011; McColl, 2012; Loye et al., 2012; Ballantyne and Stone, 2013; Ambrosi and Crosta, 2006, 2011; Leith et al, 2014). When the microclimate is similar, and in absence of a clear external trigger, the stress distribution within the rock mass and the consequent progressive damaging processes would define its resistance (Matasci et al., 2011; McColl, 2012). The stress history materializes in the rock mass geomechanical properties, primarily controlling rock slope stability and long-term equilibrium (McColl, 2012). This explains why geomechanical properties are more relevant than more readily available terrain attributes (e.g., lithology, aspect) for the understating and modelling of rock slope degrading processes (Kratutblatter and Moore 2014).

Fracture density, water circulation and weathering conditions are widely recognized as the key crucial geomechanical properties for rock slope instability (Moore et al., 2009; Ballantyne, 2002; Matasci et al., 2011; Jaboyedoff et al., 2011; McColl, 2012; Wei Wei et al., 2014; Krautblatter and Moore, 2014; Scott and Wohl, 2019; Hartmayer, 2020). These properties develop over time in a differential way, individuating weakness zones (i.e., “critical paths”, Einstein, 1983) that might be profoundly different from the surrounding rock masses (Hall et al., 2012). Geomechanical properties are the expression of processes that coexist in the rock mass system, and sometimes generate complex and generally positive feedbacks (i.e., with a mutual reinforcing mechanism - Viles, 2013; Krautblatter and Moore, 2014). Fracture density and orientation control valley morphology and rock mass wasting processes (Loye et al., 2012) and are considered not only the result of tectonic and exhumation stresses, but also of paraglacial adjustments and consequent stress release (Augustinus 1992; 1995; Ballantyne, 2002). In general, slopes with closely spaced fractures are more susceptible to rock slope instability (Scott and Wohl, 2019). Commonly, in literature, distance from fault is used as a proxy for instability processes

and is consequently included as a predictor in landslide susceptibility models (Reichenbach et al., 2018). In the case of rock mass instabilities, and particularly of rockfalls, for which the dependence from the local fracture pattern is crucial, this approach is likely to fail as the joint sets related to tectonic stresses are only one component of the rock mass system. Indeed, the presence of non-tectonic joints is equally important. They are mainly extensional fractures usually sub-parallel to the valley (Nichols, 1980), thus relevant for rock mass detachments. Their genesis is potentially triggered by cyclical phases of glacial loading and unloading, assumed that sufficient horizontal stresses had been developing during glaciations through erosion, tectonic and lock-in stresses, and were maintained during ice retreat (McColl, 2012). Rock masses are exposed to the atmosphere and therefore are subject to the combination of physical, chemical, hydrological and biological processes. These processes act non-linearly in time and may be referred as "multistage weathering", which lead to the degradation of the intact rock constituting rock bridges (De Vilder et al., 2017). In detail, "multistage weathering" acts extending the existing micro-defects, nucleating new fractures and decreasing existing fractures toughness, cohesion and frictional resistance (Ballantyne, 2002; Viles, 2013; Krautblatter and Moore, 2014). Therefore, the process triggers a feedback loop, since it provides additional surfaces and opening ways for weathering processes (McColl, 2012; Scott and Wohl, 2019). Of particular relevance in terms of fracture propagation is the weathering that takes place at the interface between a rock bridge and a discontinuity, as it is the zone where the stress concentrates most (Collins and Stock, 2016; De Vilder et al., 2017). Moreover, these weakening processes and stress redistributions may lead to variations in hydraulic conductivity and drainage patterns (Crosta et al., 2013; Wei Wei et al., 2014). Specific reasons for failure linked to water circulation could be excess pressures in joints leading to fractures propagation and opening; water freezing and thus drainage inhibition (McColl, 2012); frost cracking and frost weathering (Hales and Roering 2007; Matsuoka 2008); ice segregation-induced subcritical cracking (Draebing and Krautblatter, 2019).

As the goal was to produce a statistically-based susceptibility model, it was necessary to disentangle these complex interactions and feedbacks and synthetize them in measurable and self-standing variables. Among the several rock mass and joint metrics measurable on an outcrop, given the rationale above, the focus was on three main target properties: Joint Volumetric Count Jv (fractures/m³), Weathering index Wi (-), and Equivalent Permeability Keq (m/s), all representative of rather shallow conditions of the rock masses and in this study measured on outcrops from 10 to 150 m wide (median length of surveyed rockfall scarps: 126 m). Joint Volumetric Count represents the volumetric fracturing degree of a rock mass and was defined as the number of joints intersecting a volume of 1 m³ and is obtainable from the measured average spacing of each set of joints (Palmstrom, 1982). It is independent from the

discontinuity type, being a proxy not only of the tectonic history but also of the other stress redistribution processes acting on rock slopes at a local level, relevant for the spatial scale of the investigated instability, and contributing to define the observable fracture pattern. Weathering index represents the rock mass weathering conditions. It is calculated as the ratio between Schmidt Hammer rebounds on natural and abraded joint surfaces, following the well-known relationship between rebounds and surface age of exposure (e.g., Stahl et al., 2013; Stahl and Tye, 2020). Weathering may seem the less suitable process to be represented by a static property; however, it was quantified that a substantial amount of weathering of rock bridges can occur before rock block failure, evolving over periods long enough for weathering to take place (De Vilder, 2017), thus making Wi representative at the temporal scale of the susceptibility map. In addition, it is crucial to concentrate on the weathering predisposition of the parent material, which is strictly connected to rock mineralogy and weak planes, given similar microclimate and topography (Hall et al., 2012). The property Wi may be referred to this predisposition as it is measured on the same material, both on the natural surface and eliminating the weathered layer.

Keq was derived from the permeability tensor $\overline{K}$ introduced by Kiraly (1969) and modified by Coli et al. (2008):

$$\overline{K} = \frac{g}{12v}\sum_{i=1}^{N} f_i e^3{}_i \, [I - \overrightarrow{n_i} \times \overrightarrow{n_i}] \qquad \text{Eq. 3.1}$$

where g is gravity (9.81 m/s^2), v is the kinematic viscosity of water (3.2e^{-6} m^2/s), N is the total number of discontinuity sets, f is the average frequency of the i^{th} set of discontinuities (m^{-1}), e is the average hydraulic aperture of the i^{th} set of discontinuities (m), I is the identity matrix and $\vec{n}|n_1, n_2, n_3|$ is the dimensionless unitary vector normal to the average plane of the discontinuity set. Hydraulic aperture was defined by Barton (2004a,b), where the physical measured aperture E is corrected with JRC (Joint Roughness Coefficient) to find the hydraulic aperture e as:

$$e = \frac{E^2}{JRC^{2.5}}, \qquad \text{Eq. 3.2}$$

Specifically, once obtained the tensor $\overline{K}$, the three components of the principal diagonal were used to calculate Keq as (Guo et al., 2015):

$$K_{eq} = \sqrt[3]{k_{11}k_{22}k_{33}} \qquad \text{Eq. 3.3}$$

As Keq calculation is strongly based on joint apertures, it consequently represents the hydrogeological behaviour of the rock mass at the outcrop scale.

3.4.2. Expansion of the geomechanical dataset: field survey optimization and realization

As the spatial distribution of the pre-existing geomechanical surveys was not fully suitable for regionalization (clustering along roads), the addition of sampling points through an optimization strategy was necessary. The model-based Spatial Simulated Annealing - SSA (Van Groeningen, 1999) algorithm was adopted, adapting the steps proposed by Brus (2019). SSA is an iterative random search procedure (Brus, 2019) that optimizes a custom target function based on the pre-existing points (a comprehensive description of the method is given in Section 2.1). Due to the previously mentioned dataset heterogeneity, Jv was found to be the only target property collected at all survey points. For this reason, a preliminary regionalization of this property by ordinary kriging – already used for Jv and other geomechanical parameters in the upper part of San Giacomo Valley by Ferrari et al. (2012, 2014) – was carried out. The associated prediction error variance was used as the target function to obtain 25 new survey locations. Initial temperature *T0* is one of the most sensitive parameters in SSA to avoid local minima (Ameur, 2004). To find a suitable *T0*, the minimum temperature corresponding to the probability of acceptance of the initial iteration equal to 1 and 0.95 (Nunes et al, 2006) was computed, varying cooling rate (i.e., 0.5, 0.8 and 0.95) and comparing running time and final point locations.

Detailed geomechanical field surveys were performed at the identified rock masses (minor offsets from the exact SSA points were due to logistic choices) and some pre-existing points were contextually revisited to acquire the necessary missing properties. Due to the high mountain environment involved in the field work, in three cases the exact selected locations were unreachable: in such situation more easily reachable areas with similar characteristics in terms of geological unit and slope aspect were alternatively selected. Moreover, the multivariate environmental similarity surface (MESS; Elith et al., 2010), based on several environmental covariates (elevation, aspect, slope, longitude, latitude), was calculated both before and after the field survey, using the R package *dismo* (Hijmans, 2013). The goal was to attest if, beyond kriging variance minimization, the newly acquired information would improve the representativeness of the geomechanical dataset in a complex mountain environment. Additional details on MESS are given in Section 2.1.

3.4.3. Regionalization of geomechanical predictors

The regionalization of the geomechanical properties was necessary as spatially distributed predictors are required for susceptibility modelling. First of all, the pairwise correlations (Pearson's correlation coefficients and p-values of t-tests) among geomechanical properties and with environmental variables such as elevation, latitude, longitude, slope and aspect – expressed in terms of northness = cos(aspect) and eastness = sin(aspect) – were calculated and analysed to find preliminary patterns in the dataset.

Secondly, different deterministic (inverse distance weighting - IDW, thin plate spline - TPS), geostatistical (ordinary kriging, kriging with external drift) and regression (geographically weighted regression - GWR) regionalization techniques were evaluated. Details on the regionalization techniques can be found in Section 2.2. Each technique was applied testing either several model parameters or covariates combinations. The analyses were performed in the R environment (https://www.r-project.org/) by means of the libraries *gstat* (Pebesma, 2004), *fields* (Nychka et al., 2017) and *spgwr* (Bivand and Yu, 2021). For IDW, different powers were tested (i.e. 1.5, 2, 2.5, 3, 3.5, 4,4.5, 5). For geostatistical techniques, different variogram models were evaluated (e.g., exponential, circular, spheric, gaussian). For TPS, either coordinates alone or in combination with altitude, slope or aspect were tested as covariates. For GWR, all the possible model parameters (i.e., fixed or adaptive bandwidth, gaussian or bisquare kernel and CV or AIC method) and covariates combinations (one to all covariates) were tested.

Model performances were assessed through a leave-one-out cross-validation (LOO-CV), calculating goodness-of-fit and error metrics (correlation coefficient r of modelled and observed values, and NMRSE). For geostatistical techniques, starting from the selected variogram model on the entire dataset (i.e., "general model"), range and sill were automatically fitted on each cross validation set and for this reason they may slightly differ; to verify the consistency and the maximum variation from the general model, histogram of sill and range coming from the LOO-CV were calculated. Finally, for each property, the results coming from each technique were presented and compared. The maps considered as the most reliable based on performance indices and spatial patterns, were selected as subsequent susceptibility model geomechanical predictors in the next steps.

The regionalization was performed both on the whole study area (i.e., Global Domain) and by splitting the domains in two ways (Figure 3.4). The first split follows the topographic subdivision of the San Giacomo and Bregaglia Valleys. The second relies on geo-structural borders: the Truzzo Granite Unit and the zone where, approximately, the main structural lineaments direction varied from E-W to N-S were set as the limiting criteria between the two domains. The first domain (Domain 1) comprises the North-Central part of San Giacomo Valley where the structural lineaments are mainly directed N-S, whereas the second domain (Domain 2) comprises Bregaglia Valley and the Southern part of San Giacomo Valley, which was characterized by mainly E-W structural lineaments and included the Truzzo Granite Unit. The rationale behind this procedure was to find the optimal regionalization domain, as a trade-off between technique performance, physical-geological plausibility of properties values

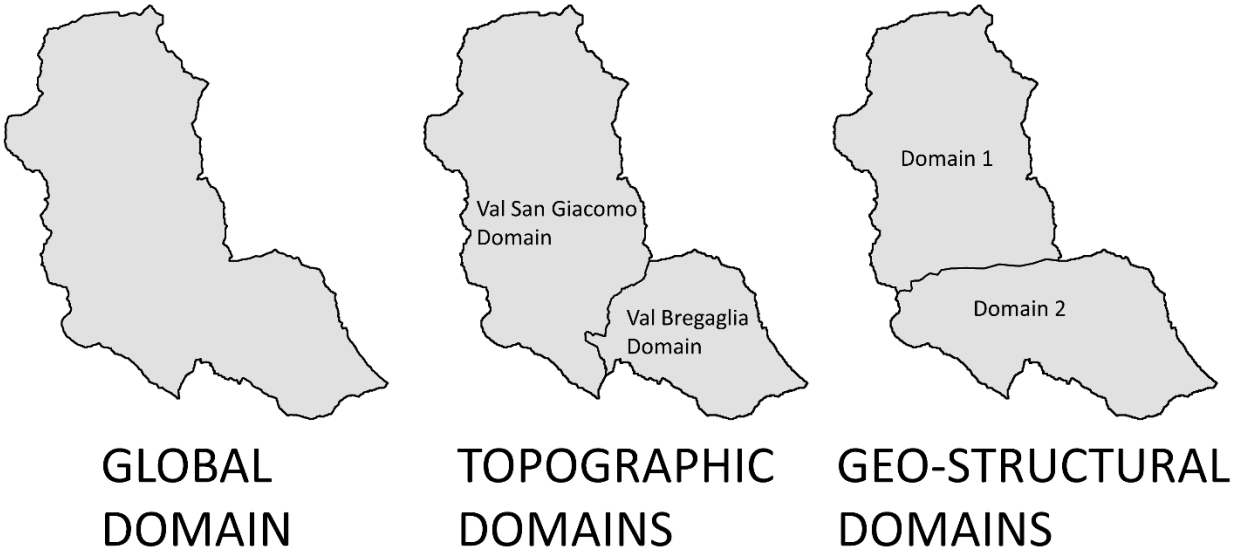


Figure 3.4 The three different regionalization domains.

distribution and, in case of subdivided domains, continuity of the spatial distribution across borders (i.e., without evident gaps and spatial artefacts at the domains contacts).

The regionalization was performed in R, at a 50 m x 50 m horizontal resolution, and a bilinear interpolation for resampling to DTM resolution to the final selected maps was applied. A minimum cut-off to the selected predicted Keq map was then added. Values lower than the typical values of intact rock permeability (according to the underlying lithology) were substituted with the intact rock values, i.e., the lowest possible value for the rock mass. The intact rock permeability values were derived from literature (Brace,1980; Kovacs, 1981; Morrow et al., 1994; Evans et al., 1997; Selvadurai et al., 2005; Sperl and Trckova, 2008; Milsch et al., 2011; Najser et al., 2011; Leclère et al., 2015; Sen et al., 2015; Duan et al., 2017; Sarout et al., 2017). All the resulting maps were reclassified using cut-offs corresponding to the recognizable flexes of the CDF (Cumulative Distribution Function) related curve (Figure 3.5).

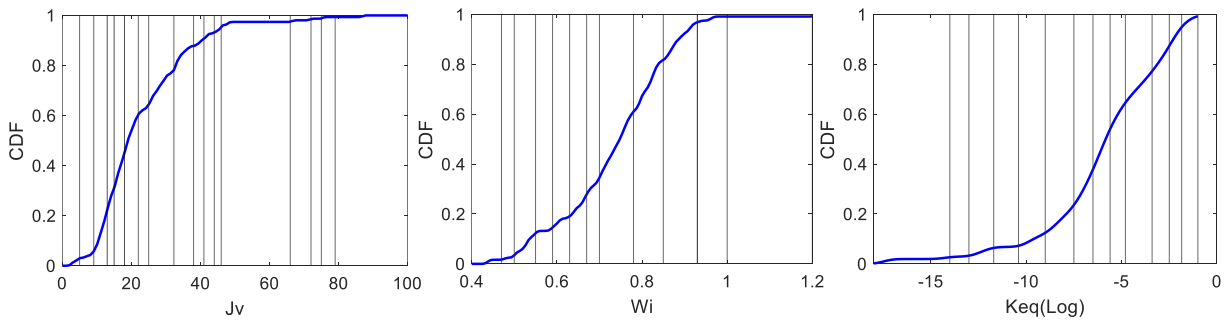


Figure 3.5 Cumulative Distribution Function related curves (blue lines) calculated for Jv, Wi and Keq datasets with the "ksdensity" function of Matlab®. The selected cut-offs (vertical black lines) were selected in correspondence of flexes of the CDF curves.

3.4.4. Rockfall inventory and geo-environmental predictors for susceptibility modelling

To uniquely identify each rockfall, despite the different sources, each event was summarized in a point feature. For the IFFI polygon-type dataset (49 events), the source point was extracted from the ROCKtheALPS dataset (https://www.arcgis.com/home/item.html?id=5304829878d04adeb4f1d1fa6fe707dc&view=list&sortOrder=desc&sortField=defaultFSOrder) and it corresponded to the highest point of the available source area polygon. For the IFFI point-type dataset (59 events), each point represented the georeferenced highest point of the crown (https://www.progettoiffi.isprambiente.it/en/methodology/). For the UNIMI dataset, 137 out of the 185 available rockfall events were associated with a mapped scarp or crown in the geomorphological map; in such cases the representative rockfall point was chosen as the bisector of the scarp. When this feature was not available (the remaining 48 out of 185), the point was placed in correspondence of the highest elevation of the mapped rockfall body (Figure 3.6). Following e.g., Rotigliano et al. 2011; Lombardo et al. 2014; Cama et al. 2015, who defined the landslide source as the point detecting site conditions responsible of previous failure, it was assumed that the rockfall deposit upper part is located at the base of the scarp originating it, or at least with similar characteristic to the unstable rock mass. This choice may introduce a possible positional inaccuracy for the 16.5% of the inventory. However, some authors, e.g., Petschko et al., (2013), representing presence either as a point from the main scarp or as a point randomly selected in a landslide body polygon, observed only small differences in the susceptibility outputs. The inventory can be classified as a geomorphological inventory (Guzzetti et al., 2012) and does not include dates of occurrence.

The binary response variable for rockfall susceptibility modelling (i.e., rockfall presence/absence) was obtained from the synthesis of these two rockfall inventory sources integrated with absence points. As absence sampling is equally crucial (Bornaetxea et al., 2018; Knevels et al., 2020), absence locations were randomly extracted from all "eligible" areas, obtained from a preliminary masking of urban areas, glaciers, water bodies and quaternary deposits (i.e., not-modellable areas) and subsequently excluding areas within an 80-m buffer from rockfall points and scarp lines. This buffer represents the average radius of the rockfall scarps surveyed in the study area. A 1:1 ratio was adopted for the extraction of absence points, following the guidelines of Hong et al. (2019), who found that this ratio is optimal when the "eligible" area approaches the 99% of the modelling domain (in the present case study the percentage was 92.3%).

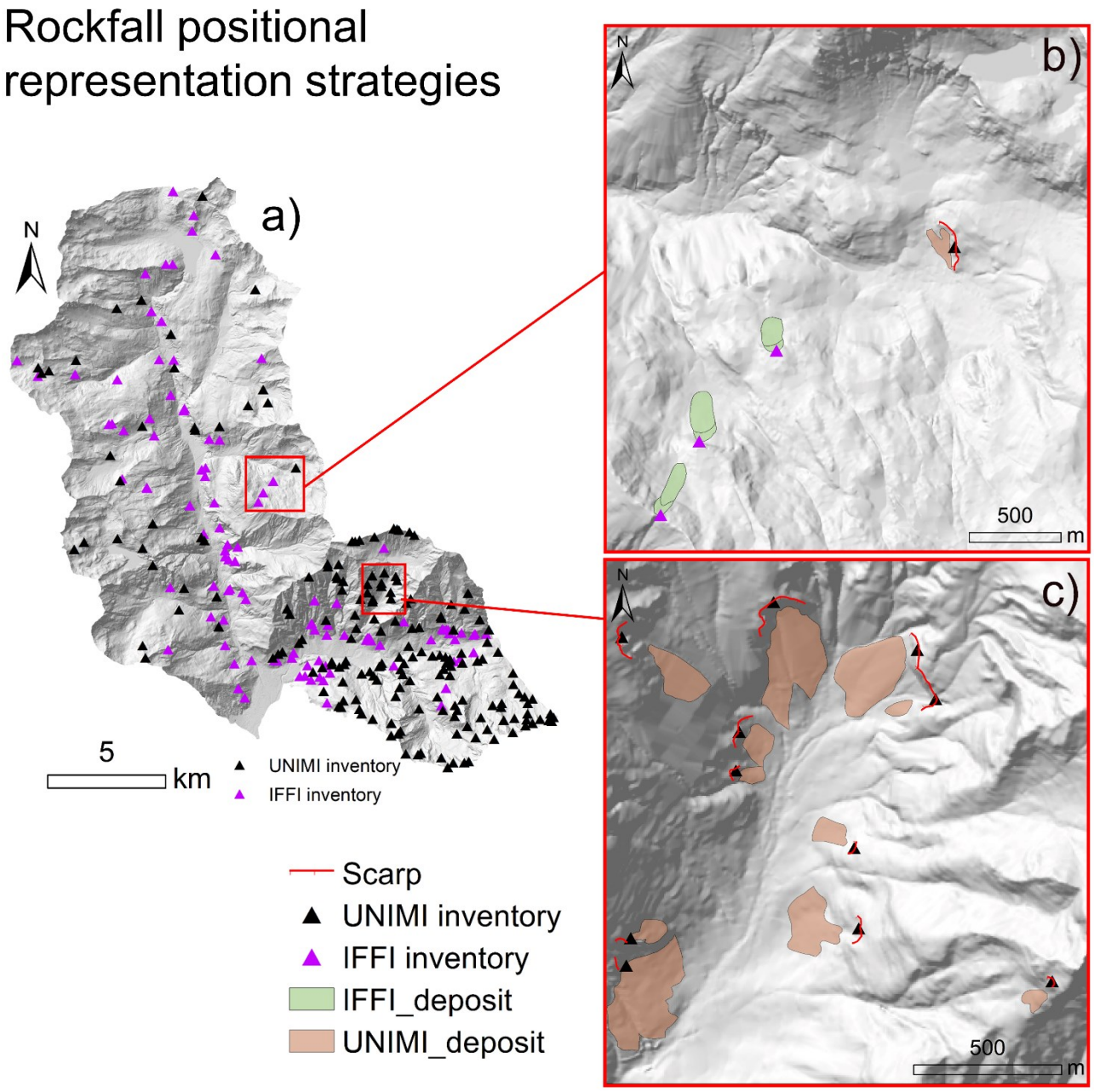


Figure 3.6 a) The integrated rockfall inventory; b) Rockfall points selected as the highest source area point for the IFFI inventory; c) rockfall points selected as the main scarp bisector for the UNIMI inventory.

Three groups of predictors were considered: topographic, geological, and geomechanical predictors. The topographic ones are the most commonly used predictors in landslide susceptibility studies (Van Westen, 2008; Reichenbach et al., 2018) as they are considered effective indicators of the slope's climatic, hydrological, and stress conditions. Elevation, slope, northness, eastness, profile curvature, plan curvature, and SAGA Topographic Wetness Index (SWI) were derived from the available DTM using the *RSAGA* package (Brenning et al., 2018). Geomechanical predictors were the previously regionalized Jv, Wi and Keq properties. Given their role in expressing rock mass geomechanical behaviour in terms of instability (see Section 3.4.1), it is expected that Jv and Wi would have a direct, and almost linear,

relationship with rockfall occurrence (i.e., more fractured and weathered rock masses are more susceptible to rockfalls). For what concerns Keq, its behaviour in the statistical model is expected to be characterized by a higher ambiguity than the other two. Indeed, high permeability may lead either to a more readily and quick dissipation of excess joint-water pressure or to high circulation of water in joints, thus influencing available moisture and chemical-physical weathering. Conversely, low permeability would mean both a more pronounced water run-off at the surface (i.e., lower infiltration) but also a rapid increase of excess joint-water pressure, favouring damage and coalescence of micro-defects, especially at the crack tips and at rock bridges surfaces. It is noteworthy to point out that these predictors represent a source of uncertainty in the model, as they were derived from the regionalization of point data. Nevertheless, the associated regionalization errors were quantified with the aim of communicating them to potential users of the final product.

To consider the hydrogeological component on rock mass instability at different spatial scale, the infiltration density predictor was introduced too. It was defined as the density of geomorphological-structural elements prone to infiltration, such as regional lineaments, trenches and counterscarps, and sinkholes. This information was derived from the detailed structural-geomorphological map (1:10000 scale) of the area and calculated on a 5m x 5m pixel basis to be consistent with the resolution of the other predictors. This predictor may be interpreted as a proxy for the infiltration and draining behaviour of slopes, linked to the broad environmental context in which rock cliffs are located. Indeed, lineaments and morpho-structures can be interpreted as weaknesses where geomorphic and gravitational processes are preferentially initiated and propagated (Selby, 1982; Cruden, 2003; Loye et al., 2012). Moreover, areas characterized by channels and concentration of flow and infiltration may localize mass wasting processes that, at the local scale, enhance joint weathering and water supply to the rock mass (Walter et al., 2012; Wei Wei et al., 2014; Scott and Wohl, 2019). Topographic predictors and infiltration density are shown in Figure 3.7.

Geology was introduced as a categorical predictor with five classes based on the geological map available for the study area (Section 3.3.1). The distinction adopted followed a lithological criterion, which is the most used in landslide susceptibility literature, as different lithologies have different ranges of strength (Catani et al. 2005; Segoni et al. 2018; 2020). The lithologies recognized in the study area include: (i) Paragneiss, (ii) Granites, (iii) Orthogneiss (iv), Shists and metasedimentary lithologies, (v) Ultramafic rocks. The class adopted as the modelling reference class was the most abundant in the study area, i.e., the paragneiss lithology. As pointed out by Segoni et al. (2020) the drawback in using this distinction may be that lithological homogenous rock-masses could differ in terms of weathering and

structures. Therefore, they proposed to use multiple and independent classifications of the geological component as models' predictors (i.e., lithologic, chronologic, structural, paleogeographic, and genetic units). In this study, since geomechanical properties, i.e., Jv, Wi and Keq, were introduced as independent predictors to enhance the spatial differentiation of rock masses at the local scale, a single lithological classification based on the protolith seemed appropriate as a descriptor of the intact rock (or rock bridges) properties.

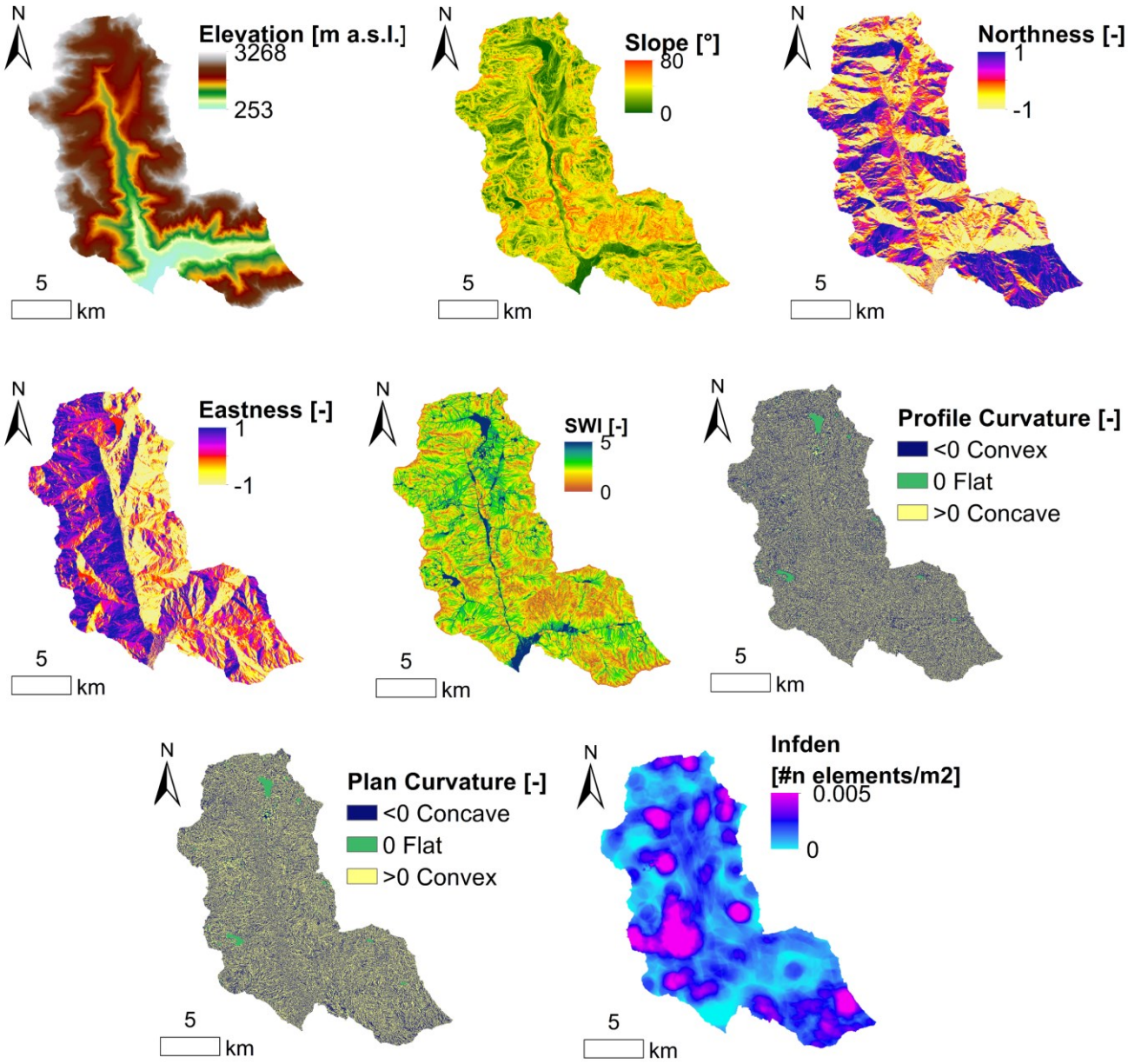


Figure 3.7 Topographic predictors and infiltration density predictor.

3.4.5. Rockfall susceptibility modelling with Generalized Additive Models

3.4.5.1 Model setup and performance assessment

Rockfall susceptibility was performed using Generalized Additive Models (GAMs), able to represent non-linear predictor-response relationships while maintaining interpretability through their additive structure (Goetz et al., 2011, 2015). Interpretability was necessary to assess the behaviour and plausibility of geomechanical predictors. Analyses were performed in R using the *mgcv* package (Wood, 2017).

Three models were built using different groups of predictors: a topographic model (TOPO), a topographic-geomechanical model (GM), and a topographic-geomechanical-geological model (GM+GEO). To achieve a good trade-off between model size and fit, variable selection through shrinkagewas carried out. Shrinkage allows penalizing out of the model predictors with low or no influence. Topographic predictors that were removed from the TOPO model were not further considered in the successive GM and GM+GEO models. The relationships between rockfall occurrence and predictors were analysed through the associated smoothing functions by means of CSF (Component Smoothing Function) plots and odds ratios for geology categories. Odds ratios represent the chance of an outcome, given a certain class in comparison to a reference one. A class with an odds ratio<1 represents a lower chance of a modelled outcome (rockfalls) in comparison to the reference class, an odds ratio>1 a higher chance, and an odds ratio around 1 means no relationship (Knevels et al., 2020).

The performance of the three models was assessed and compared through a spatial k-fold cross-validation (sCV) and the estimation of the area under the receiver operating characteristic curve (AUROC), a widely applied performance measure for landslide susceptibility models (Corominas et al., 2014; Reichenbach et al., 2018). Spatial cross-validation is preferable when dealing with spatial data, which are often subject to spatial autocorrelation (Brenning, 2005, 2012b). A spatial partitioning based on k-means clustering of coordinates was applied to derive $k=5$ folds and, to obtain results that are independent of a particular partitioning, the procedure was repeated r=100 times. Spatial CV was implemented using the R package *sperrorest* (Brenning, 2012a). As a measure of spatial transferability, the interquartile range (IQR) of the test-set AUROC values (Petschko et al., 2014) was analysed. Furthermore, the predictors' CSF from each cross-validation run were compared with the corresponding CSF obtained on the entire dataset, to assess coherency and robustness of their behaviour. For the GM+GEO model, the variability of geology odds ratios estimated in the different CV runs was assessed too. To investigate the importance of predictors, the penalization frequency coming from the application of the shrinkage option (i.e., percentage of CV runs in which the effective degrees of freedom, i.e., *edf*,

much lower than 1, here we used a 0.7 threshold) and the mean decrease in deviance explained (mDD%), calculated as in Knevels et al. (2020), were combined. Moreover, concurvity between the smoothers, i.e., the generalization of multicollinearity to non-parametric functions was calculated; high concurvity values may lead to poor and unstable parameter estimation (Amodio et al., 2014).

3.4.5.2 Susceptibility map and geomorphological plausibility

The three output maps were reclassified into five susceptibility classes (0.0-0.3 "very low", 0.3-0.5 "low", 0.5-0.7 "medium", 0.7-0.9 "high", 0.9-1.0 "very high"). Following previous studies (e.g., Sterlacchini et al., 2011; Goetz et al., 2015; Steger et al., 2016a), spatial patterns in susceptibility maps in terms of geomorphological plausibility were critically analysed. Two variation maps were produced, resulting from the difference between the TOPO and GM original susceptibility maps (i.e., before reclassification) and between the GM+GEO and GM susceptibility maps, respectively. The variation maps allowed the identification of areas where susceptibility values changed due to a modification in the predictor set and, by qualitatively comparing them with the geomorphological and geo-structural features of the area, whether they are physically plausible and coherent.

To explore the physical plausibility of the derived maps, a comparison with Sentinel-1 SAR data (PS/DS mean annual velocities) was carried out. Although these types of data are mainly suitable for the investigation of slow deformations, in particular deep-seated gravitational slope deformation (DSGSD; Frattini et al., 2018), they could also be useful to study more rapid phenomena such as rockfalls. This is feasible if the slow deformations are considered as proxies for various processes relevant for slope stability (e.g., hydrogeological circulation, creep, neotectonics). In addition, active DSGSD are usually linked to the nucleation of secondary, shallower instabilities (Crippa et al., 2020), and the presence of DSGSD was even implemented as predictor for secondary landslides (Carrara et al., 1991).

3.4.5.3 Uncertainties related to the inventory

As the rockfall inventory integrated two different sources, it was possible to evaluate the effects of potential biases in the official inventory. Indeed, the IFFI inventory (see Section 3.3.2) was compiled starting from administration reports and multi-temporal aerial photographs, therefore it could be affected by an underrepresentation of rockfalls in remote areas (Guzzetti et al., 1999; Steger et al., 2016a) and by an obliteration of geomorphic features in forested areas (Steger et al., 2017). Model performance, the behaviour of predictors, and predicted susceptibility patterns of the GM and GM+GEO models using the IFFI inventory as the training set and the UNIMI inventory as an independent test set were calculated

and analysed. The two IFFI-based models were cross validated as in Section 3.4.5.1 to assess model performance as well.

3.4.6. Rockfall susceptibility modelling with Random Forest

3.4.6.1 Model setup and performance assessment

Random Forest (RF) is another frequently used algorithm for landslide susceptibility, resulting in very high performing models (e.g., Brenning 2005; Catani et al., 2013; Paudel and Oguchi 2014; Segoni et al., 2015, 2020; Youssef et al., 2016; Lagomarsino et al., 2017; Trigila et al. 2013) thanks to its ability to manage complex interactions and collinearity issues. Moreover, it is a common practice to apply several modelling algorithms to the same case study to compare results in terms of quantitative performance (e.g., Pham et al., 2019,2020, Althuwaynee et al., 2014, Abedini, 2019, Chen W. et al., 2019,2020; Pourghasemi and Rahmati, 2018), interpretability and map user perception (Goetz et al., 2015), and geomorphological plausibility (Steger et al., 2016b). Also, the combination of models is applied to reduce the uncertainty related to the predicted susceptibility by producing ensembles (e.g., Rossi et al., 2010; Di Napoli et al., 2020; Chen W. et al., 2018; Youssef et al., 2015; Choi et al., 2012; Andan et al., 2020; Rossi and Reichenbach, 2016). With these objectives, RF was applied to model rockfall susceptibility in the study area, for sake of comparison with the GAM models in terms of predictors behaviour interpretability, quantitative performance, geomorphological plausibility, and to reduce uncertainty (or, in other words, to reveal areas of discordance and agreement between the two model outputs in a risk management perspective).

As the GM and GM+GEO models derived applying GAMs only carried on the topographic variables not excluded by shrinkage, it was deemed appropriate to develop a totally independent RF model, without preconceptions linked to previous results on which variables have to be included. For this reason, a single comprehensive model was tested, which included all the predictor variables: elevation, slope, northness, eastness, profile curvature, plan curvature, SWI, geomechanical predictors, infiltration density and geology as a categorical predictor.

Analyses were performed in R using the *ranger* package (Wright and Ziegler, 2017) for RF implementation and, combined with the *sperrorest* package (Brenning, 2012a), for hyperparameter tuning (refer to Chapter 2, Section 2.3.5). The selection of the optimal hyperparameters set was based on the optimization of two different performance measures: OOB error and AUROC, the latter derived from a spatial 5-fold cross-validation repeated r=100 times, as previously done for the GAM models. The final optimal hyperparameter set would be a trade-off between the two. The tuning procedure

adopted was the grid search, starting from a list of discrete values, following the suggested strategy of Bohemke and Greenwell (2019) for the selection of each hyperparameter set (Table 3.1), for a total of 750 possible combinations. Regarding the number of trees, a preliminary analysis of the OOB error stabilization was carried out to find the maximum *ntree* to be tested.

Table 3.1 Hyperparameter values tested during the grid search optimization procedure.

Hyperparameter	Tested values for the grid search optimization
mtry	3,5,7,9,11
Node size	2,4,6,8,10
Sample Size	0.632 (ranger default), 0.7, 0.8
ntree	100 to 1000 with step=100

3.4.6.2 SHAP values for model interpretation

Despite the usually high performing models deriving from the application of RF, some researchers expressed concerns regarding both predictors behaviour interpretability and user-friendly usability and readability of the derived susceptibility maps (e.g., Brenning, 2005, Brenning, 2012, Goetz et al., 2015). To overcome these issues, the relationship between rockfall occurrence and predictors was assessed through the application of the SHAP framework, by means of the R packages *treeshap* (Komisarczyk et al., 2021) and *SHAPforxgboost* (Liu and Just, 2021). Both packages allowed to calculate SHAP for each individual prediction of the response variable and to combine them into global explanation. Explanation plots (examples from literature in Figure 3.8) include (Molnar, 2019; Lundberg et al., 2020):

- **Feature importance plot**: it represents the SHAP-based feature importance coming from the SHAP value of each predictor averaged on all observations.
- **Summary plot**: it combines feature importance with feature effects. Each point is an observation Shapley value for a predictor. The position on the y-axis is determined by the predictor importance and on the x-axis by the relative Shapley value. The colour represents the value of the predictor from low to high.
- **Force plot**: This plot stacks the SHAP values for each observation, with different colours depending on the predictor, showing how the final prediction was obtained as the sum of

predictors' SHAPs. This plot is useful to reveal possible exceptions to the global SHAP behaviour of predictors.

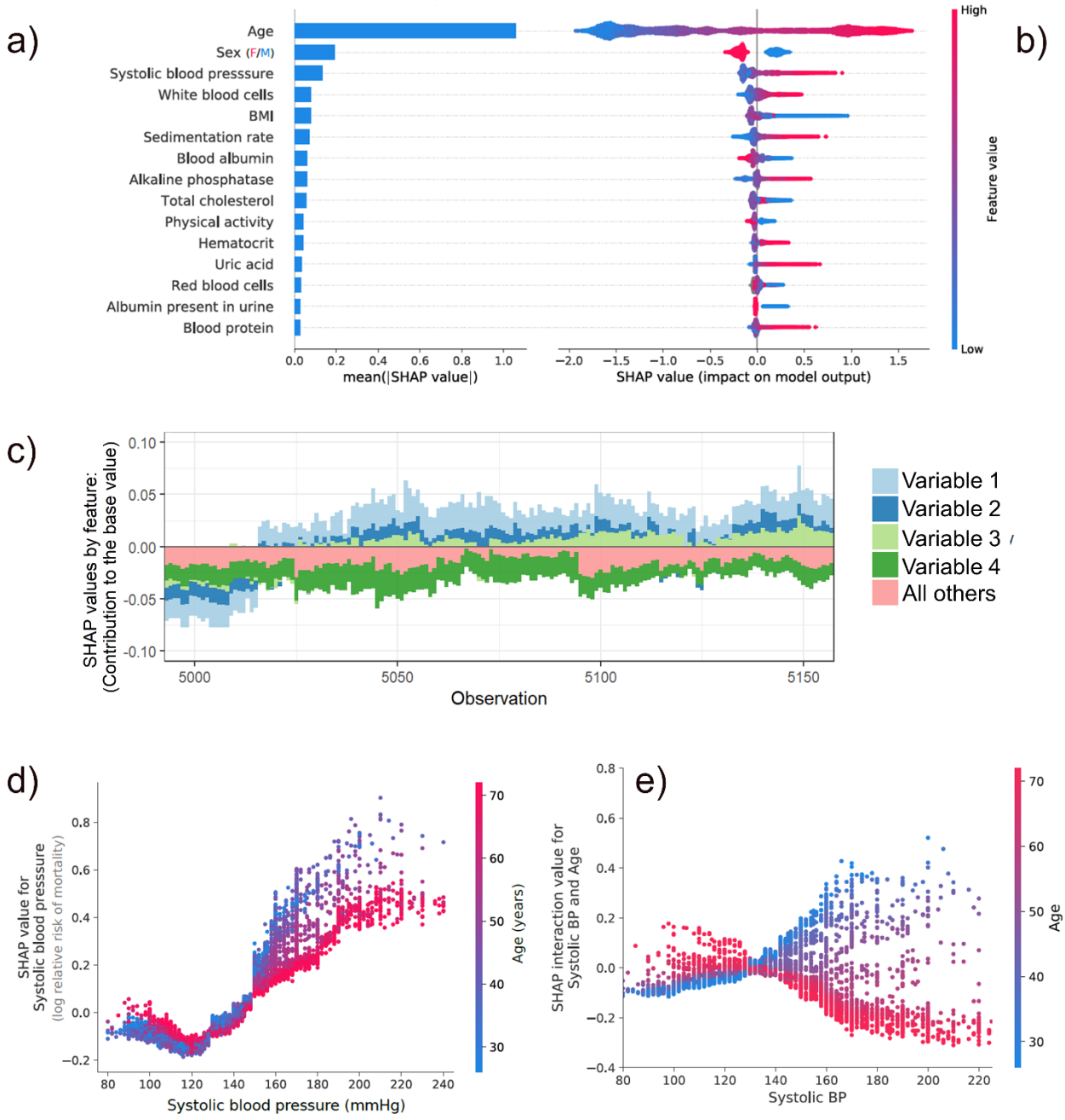


Figure 3.8 Figures in panel a,b,d,e were extracted from the study of Lundberg et al., 2019 (arXiv:1905.04610). Their study dealt with the impact of a set of health conditions on mortality in an US population sample. a) feature importance plot showing age and sex having the highest impact on mortality. b) summary plot showing both the importance of variables on mortality, and their behaviour. For example, old people, with high systolic blood pressure (BP) are linked to an increased mortality. c) the force plot example was extracted from Liu and Just (2020), authors of the package SHAPforxgboost: (https://github.com/liuyanguu/SHAPforxgboost/). The SHAP force plot stacks the SHAP values for each observation showing how the final prediction was obtained as the sum of predictors' SHAPs. d) dependence plot of the predictor systolic blood pressure, showing its contribution to mortality in Lundberg et al., 2019 (arXiv:1905.04610). These plots may be also coloured with a second predictor, in this case age- otherwise predictors may be viewed singularly. e) Interaction effect plot between systolic blood pressure and age. A high systolic BP is more concerning in terms of mortality in the young population.

- **Dependence plot**: It plots the SHAP values against the observation values for each variable. This plot could be interpretable as the analogue of the CSF plots in the GAM modelling framework. A second variable could be added as the points colour to interpret possible correlation between predictors.
- **Interaction effects plot**: This plot separates the impact of a variable into main effects (equal to the dependence plot) and pure interaction effects. It could be interpreted as the difference between the SHAP values for feature *i* when feature *j* is present and the SHAP values for feature *i* when feature *j* is absent.

3.4.6.3 Uncertainties related to the model selection: comparing GAM and RF

All the global explanation plots mentioned in the previous Section will be discussed in terms of physical plausibility of predictors behaviour and compared to the CSF plots derived from the GAM models. Besides predictors behaviour, the comparison between the GM+GEO models derived from GAM and RF was carried out at the level of final susceptibility maps and performance. AUROC and IQR of both models were compared to declare the best performing and transferable one, at least at the quantitative level. The RF derived output map was reclassified into the same five susceptibility classes adopted for the GAM based maps (i.e., 0.0-0.3 “very low”, 0.3-0.5 “low”, 0.5-0.7 “medium”, 0.7-0.9 “high”, 0.9-1.0 “very high”). Susceptibility maps were compared in terms of susceptibility difference and susceptibility class change. Moreover, adopting the strategy from Di Napoli et al. (2020), the technique of the Averaging Committee with two different susceptibility cut-offs, namely 0.5 (i.e., more conservative) and 0.7 (i.e., more suitable for prioritize intervention) was applied. Firstly, the GAM and the RF derived susceptibility maps were transformed into binary data (e.g., for the 0.5 cut-off, a value of 0 was attributed to susceptibility values<0.5 and a value of 1 was attributed to susceptibility values >0.5); in this way, depending on the cut-off, each model votes for the landslides being either present or absent. This measure gives both a prediction and a measure of uncertainty. When the prediction is 0 or 1, it means that all the models agree to predict low or high susceptibility respectively; when the prediction is 0.5, it means that the two models disagree in predicting susceptibility.

3.4.7. Combining susceptibility and SAR with operational purposes

False negative errors in landslide susceptibility modelling (i.e., unstable terrains misclassified as stable) could be related, in a risk management perspective, to high social and economic costs, as those areas may be incorrectly used without restrictions (Frattini et al., 2010; Ciampalini et al., 2016). In an operational perspective, Ciampalini et al. (2016) proposed a procedure to integrate landslide susceptibility modelling results and SAR data in terms of mean annual PS/DS velocity, to update and

increase the reliability of the susceptibility map, by means of an empirical contingency matrix. These data are especially suitable for slow or very slow-moving landslides (Ferretti et al., 2000, 2001), theoretically preventing the application of the integration strategy of Ciampalini et al. (2016) to analyse fast phenomena as rockfalls. Nevertheless, active slope deformations, in complex alpine contexts, are often associated with rock mass related processes such as slope hydrogeological circulation and drainage, neotectonics, progressive deformation and fracturing, which indirectly individuate active contexts suitable for rockfalls occurrence. More specifically, the presence and activity state of morpho-structural features associated to slow slope deformations play a crucial role in the dissection of the rock masses and are mainly responsible of the changes in the hydro-mechanical properties of the slope system (Crosta et al., 2013). Moreover, the morphological response types of paraglacial rock slopes (i.e., large-scale catastrophic failure, slow and progressive rock mass deformations and periodic small-scale rockfalls) are not mutually exclusive, making the limit between these categories "blurred" (Ballantyne, 2002). Assuming these dynamic relationships between slow deformations and rockfalls, the use of SAR products for rapid phenomena might induce both an underestimation and overestimation of rockfall activity, depending on the characteristics of the area. An underestimation may be expected as rockfall occurrence is not necessary linked only to DSGSD; an overestimation, although in favour of safety, may be expected in DSGSD active contexts, as rockfalls are not always necessary present, even though a very common feature. With these possible mismatches in mind, the procedure of Ciampalini et al. (2016) was applied to the Valchiavenna case study. The objectives were to: (i) integrate the Mean Ensemble rockfall susceptibility map (i.e., the average between the GAM and RF output maps) by means of a contingency matrix including susceptibility classes and SAR velocities; (ii) give an operative connotation to the Average Committee map; (iii) verify if the procedure is applicable to rapid phenomena by considering an active context as a proxy for rockfall occurrence; (iv) anchor the susceptibility map to a precise time-frame, resulting in potential updates at different temporal scales.

The procedure, which follows the steps proposed by Notti et al. (2014), updated by Ciampalini et al. (2016) and adapted to the present case study, is based on the Sentinel-1 derived SAR data available both in ascending and descending geometries, from October 2014 to February 2021. The procedure was implemented in the ArcGIS® 10.2.2 software and is summarized below:

Step 1: Slope (*S*) and aspect (*A*) were derived from the available DTM (5 m x 5 m) and transformed in radians.

Step 2: V_{LOS} (i.e., the average annual velocity of PS and DS in the satellite LOS direction), both related to the ascending and the descending geometries, were transformed from point data to raster layers, averaging the values in cells 100 m x 100 m. The resulting raster where then resampled by a bilinear interpolation to match the susceptibility map resolution.

Step 3: The C coefficient (Colesanti and Wasowski, 2006; Plank et al.,2012; Notti et al., 2014), representing the suitability of the SAR sensor geometry to record a slope movement, was calculated as:

$$C = (nlos \cos(S)\sin(A - 1.571)) + elos(-1\cos(S)\cos(A - 1.571)) + (hlos \sin(S))) \quad \text{Eq. 3.4}$$

Where *nlos, hlos* and elos are the direction cosines of the LOS, obtainable from the incidence angle θ and LOS azimuth δ in radians, thus exclusively depending on the satellite sensors orbit (see Figure 3.9 and Table 3.2 for details). Positive values of C indicate the suitability of the SAR sensor geometry to capture the landslide movement.

Step 4: V_{LOS} velocity was projected in the direction of the steepest slope to obtain a Vslope layer for each acquisition geometry (i.e., ascending or descending). Indeed, the most reliable displacement can be measured in the direction of the local maximum slope gradient, which could be considered as the most probable direction for a slope movement (Cascini et al., 2010; Notti et al., 2010; Bianchini et al., 2013; Herrera et al., 2013; Ciampalini et al., 2016). For this reason, the Vslope parameter is especially suitable for phenomena where the landslide principal movement is parallel to the slope. Rockfalls, before free falling, are usually linked to the dislocation along traction planes, parallel to the slope. Vslope is represented by the ratio between V_{LOS} and the coefficient C. To reduce the exaggeration of the projections when C tends towards 0, the correction proposed by Notti et al. (2014) was applied: C = –0.2 when –0.2 < C < 0 and C = 0.2 when 0 < C < 0.2. Moreover, when Vslope>0, the raster pixel of the

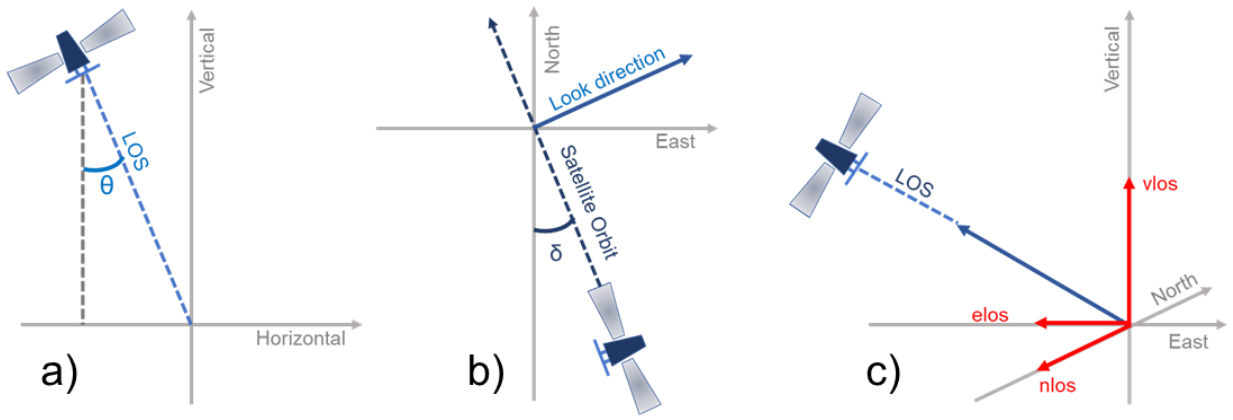


Figure 3.9 Details on satellite orbit configuration: a) incidence angle; b) LOS azimuth; c) nlos, hlos and vlos direction cosines.

resulting map was set null. Reasons are that Vslope can be positive either if the SAR sensor geometry is suitable to catch the slope movements (i.e., C>0) but V_{LOS} is positive (i.e., the landslide is going up the slope, which is a very improbable condition), or when the SAR sensor geometry is not suitable to catch the slope movement (i.e., C<0) although V_{LOS} is negative Notti et al. (2014).

Step 5: the ascending and descending data sets were combined by merging the Vslope rasters calculating their average value.

Step 6: the final Vslope layer was masked in the same way as the susceptibility map, obtaining the values only on rock outcrops (i.e., the "eligible area" in Section 3.3.4).

Table 3.2 Details on values of director cosines, incidence angle and LOS azimuth for each geometry and track.

Geometry and track	nlos	hlos	elos	Incidence angle	LOS azimuth
Ascending T15	-0.116	0.759	-0.641	40.64°	10.23°
Descending East	-0.102	0.749	0.655	41.49°	8.89°
Descending West	-0.108	0.799	0.591	36.94°	10.38°

The final Vslope layer was then combined in two ways with two different ensemble susceptibility maps, deriving from the integration of the GAM and RF results. The first product was named SAR Integrated Susceptibility Map, which was obtained following Ciampalini et al. (2016), through the integration and verification of the classified Mean Ensemble rockfall susceptibility map, by means of an empirical contingency matrix (Figure 3.10) based on Vslope values. The aim of this map was to increase the susceptibility degree of active areas, while maintaining unmodified the cells already characterized by high susceptibility degree. Indeed, the absence of active movements does not necessary imply not susceptible areas, as the SAR data are related to a particular time frame. To create the contingency matrix, Vslope values were reclassified into five intervals (i.e., the same number of susceptibility classes), which were based on multiples of the standard deviation of Vslope (σ=3.7 mm/year).

	Contingency Matrix	Vslope (mm/y)				
		0-3.7 (0-1σ)	3.7-11.1 (1σ-3σ)	11.1-18.5 (3σ-5σ)	18.5-25.9 (5σ-7σ)	>25.9 (>7σ)
Susceptibility	Very Low	0	1	2	3	4
	Low	0	0	1	2	3
	Medium	0	0	0	1	2
	High	0	0	0	0	1
	Very High	0	0	0	0	0

Figure 3.10 The empirical contingency matrix used to combine susceptibility classes and Vslope. Numerical values represent the additional number of susceptibility classes given to each pixel, depending on the combination between the Mean Ensemble Susceptibility map and Vslope.

The second product was named Intervention Priority map and was derived as the integration between the Averaging Committee Ensemble map (0.5 susceptibility cut-off) with the Vslope values, by means of an Intervention Index matrix (Figure 3.11). For this second matrix, Vslope was reclassified with the same five intervals used to derive the SAR Integrated Susceptibility Map. The rationale behind the Intervention Priority map was to furnish practical suggestions in managing the uncertainties related to the model ensemble agreement (i.e., 1 or 0 classes) and disagreement (i.e., 0.5 class), by updating the map with slope activity evidence.

For instance, in those areas where the GAM and RF showed a disagreement (i.e., high ensemble uncertainty), but active deformations are recorded, local analysis to deepen the knowledge of the area and causes of activity may be required. Conversely, in areas where the two models showed agreement in classifying the location as highly susceptible (i.e., class 1) and active deformation are recorded, monitoring measures of the slopes are recommended.

Intervention Index		Vslope (mm/y)				
		0-3.7 (0-1σ)	3.7-11.1 (1σ-3σ)	11.1-18.5 (3σ-5σ)	18.5-25.9 (5σ-7σ)	>25.9 (>7σ)
Average Commitee	0	NI		CE		
	0.5	NI		HP-2		
	1	MP		HP-1		

Legend

NI = No Intervention required.
CE = Check of local processes responsible of detected movement or possible errors in the LSM.
MP = Models agree in High Susceptibility, no actual detected movements. Careful observation of the evolution of the slope is recommended.
HP-1 = Models agree in High Susceptibility + detected active deformation. Monitoring measures recommended.
HP-2 = Models disagree in High Susceptibility + detected active deformation. Local analysis to solve uncertainty are recommended.

Figure 3.11 The intervention index matrix used to combine the Averaging Committee ensemble and Vslope.

3.5. Results

3.5.1. Geomechanical properties: sampling and regionalization

The preliminary results of the Jv regionalization, performed by means of ordinary kriging, are shown in Figure 3.12a and Figure 3.12b and were used as input data for the optimal selection of new survey points. The variography related to ordinary kriging performed on Jv before dataset updating resulted in an anisotropic variogram of gaussian type with a lag equal to 1000 m and a maximum range of 11929 m in the SW-NE direction, approximately parallel to the main shistosity dip direction, which denoted the main orogenic stress direction. The associated kriging variance was used as the minimization function for the SSA algorithm. The sensitivity analysis on T0 resulted in a T0=0.1 and T0=0.002 for a probability of acceptance of the initial iteration equal to 1 and 0.95, respectively. However, as small differences in terms of point locations and running time (probably related to the small size of sampling, i.e., 25 points) were observed, the configuration linked to the highest initial temperature (i.e., the most conservative choice, with both slightly higher running time and precision) was preferred. Figure 3.12c shows the SSA-selected points and the survey locations consequently used in addition to the old dataset.

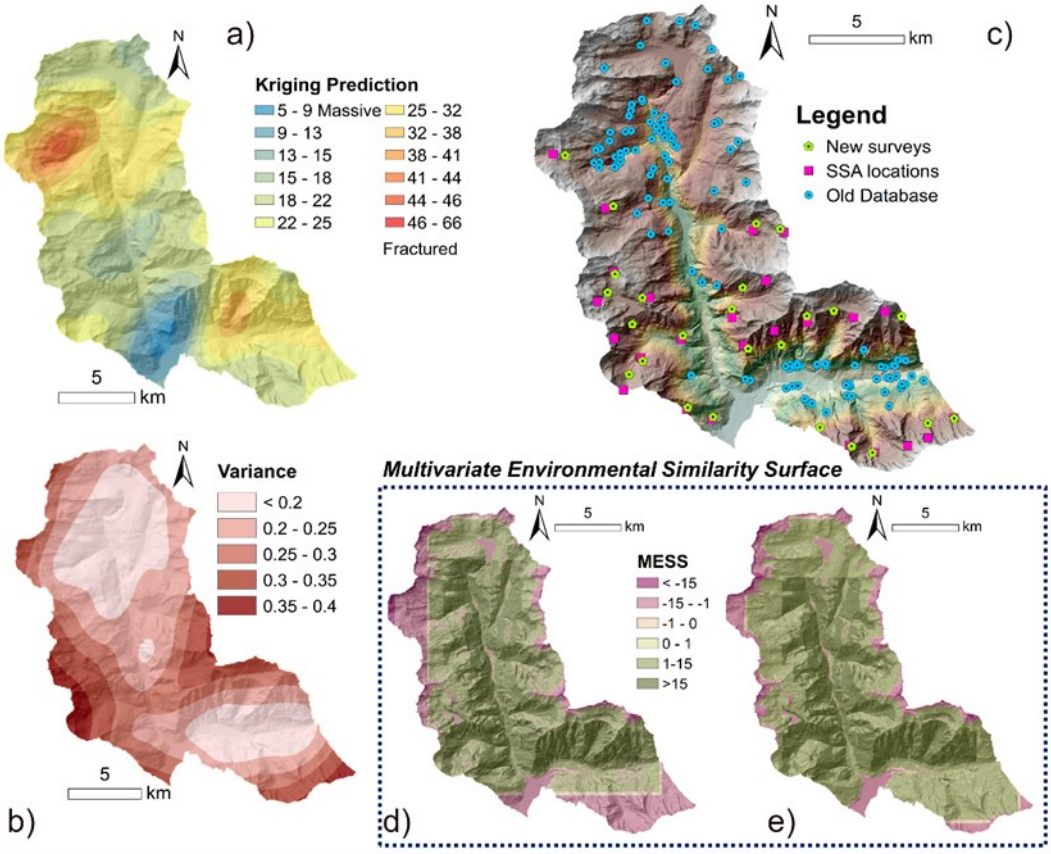


Figure 3.12 a) Jv prediction by ordinary kriging pre-field survey; b) Jv associated kriging variance pre-field survey; c) map showing old database locations, SSA selected locations, new survey locations; d) and e) MESS values before and after field survey.

Regarding the MESS resulting maps, the updated inventory enhanced the similarity and the representativeness of the geomechanical dataset, particularly on the southern slope of Bregaglia Valley and the upper western slope of San Giacomo Valley (Figure 3.12d and Figure 3.12e). The addition of the new points limited negative values of MESS to 15.9% of the study area in comparison to the 26.2% associated to the original dataset. The new surveyed points allowed the geomechanical dataset to be updated and homogenized for the regionalization phase.

After field survey completion, the procedure consisted of regionalization of the three selected target properties, both in the global and in the split domains. The results on the divided domains were quite similar to the whole domain regionalization, both in terms of model performance and spatial pattern, even if in some cases (but alternatively, only for one of the two domains) they resulted in a slightly better performance than for the Global Domain. However, with the domains subdivision, some inconsistencies in values continuity at the domain borders were frequently observed. For the topographic subdivision, the main difference at the domain borders was represented by the slope aspect and consequently to possible insulation and thermal differences, which may explain differences in weathering but not in joint frequency and aperture. For the geological domain subdivision, contrasts at the borders are more physically plausible as may explain differential strength, fracturing and weathering behaviour. However, the continuity of the structural regional lineaments across the two domains and the juxtaposition, at the domains' borders, of granitoids (Truzzo Granite Unit) and gneissic bodies of the Tambò basement hampers the rising of extremely contrasting values, more likely attributable to interpolation-linked artefacts rather than to real physical contrasts. Therefore, the Global Domain regionalization was preferred for three reasons. First, since the spatial pattern is mainly preserved amongst the subdivided domains and the Global Domain, with only slight performance and error oscillations, the Global Domain may be considered as the optimal domain for regionalization of the geomechanical target properties. Second, even with their peculiarities, which led to both a topographic and a geological domains delineation, the two orthogonal valleys share a common geological, structural, and geomorphological history and evolution, thus a Global Domain is still a reasonable and physically sound choice. Last, the Global Domain regionalization was carried out on a higher number of data than for the subdivided domains, thus being more statistically representative of the analysed population than the sub-domains. Following, performance metrics and best maps for each property and for each regionalization domain are presented giving more emphasis to the results of the Global Domain.

For Jv, ordinary kriging outperformed the other techniques in the Global Domain (Figure 3.13a,b and Table 3.3). The variogram (Figure 3.13c) was best fit by an anisotropic exponential model with a nugget

effect of 0.14, a partial sill of 0.41, a maximum range of 12 km in the NW-SE direction, and a selected lag of 1000 m. LOO-CV correlation between observations and predictions increased from 0.43 (pre-field survey) to 0.49 (post-field survey), while NMRSE decreased from 14.1% to 13.7%. The range and sill deviations from the ones of the general model remained quite low during the LOO-CV, confirming the robustness of the selected variogram model and parameters (Figure 3.13d). Globally, the associated variance decreased too, mainly due to an increase performance at high altitudes and in the Bregaglia Valley. In comparison with the pre-field kriging variance (Figure 3.12b), a slight increase can be noticed only locally, in the NE part of the San Giacomo Valley (Figure 3.13b).

These results demonstrated that the adopted sampling strategy was successful in reducing the uncertainty associated to this predictor. The change in the maximum range direction is not surprising since new spatial patterns could be revealed introducing new data. Indeed, in most cases, the model adopted for sampling and the model for subsequent statistical inference (i.e., regionalization) differ, as the new

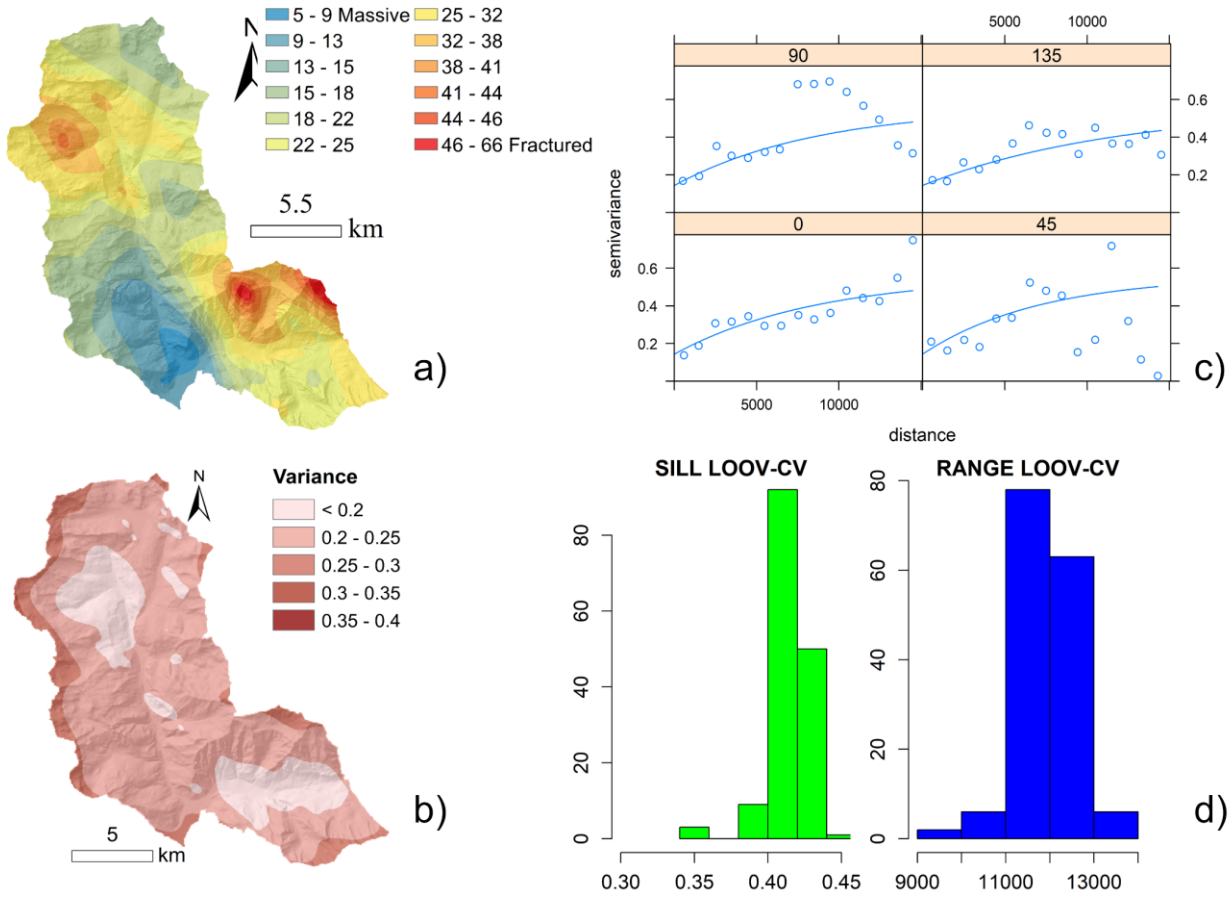


Figure 3.13 a) Jv map derived from the anisotropic ordinary kriging applied on the Global Domain and b) associated variance; c) experimental variograms of Jv (points) and variogram models (lines) in the 0,45,90 and 135 degrees directions; d) histograms representing the variations of Sill and Range coming from the different LOOV-CV sets. Having Jv a lognormal distribution, semivariance values in in c) and sill values in d) are in lognormal scale.

sampled data are used to update the postulated preliminary model used for the sampling design (Brus, 2019). Moreover, the new anisotropy direction was the same reported for Jv data limited to the upper San Giacomo Valley by Ferrari et al. (2012) and it is geologically meaningful as it is sub-parallel to the Forcola fault system, and to the tensional joints resulting from glaciers retreat in the San Giacomo Valley.

Table 3.3 Regionalization model parametrization, performance and error metrics results for Jv in the Global Domain.

Technique	**Parametrization**	**r coefficient**	**NMRSE**
IDW	Idp=3	0.28	0.165
TPS	Covariate=coordinates+elevation	0.37	0.148
GWR	Bisquare kernel, adaptive bandwitdth, AIC criterion Covariate=Latitude	0.27	0.150
GWR+IDW on residuals	Bisquare kernel, adaptive bandwitdth, AIC criterion Covariate=Latitude Idp=2	0.31	0.157
OK	Exponential Isotropic Lag=1000 m	0.48	0.138
OK	Exponential Directional 135° Lag=1000 m	0.49	0.137

Regarding the results on the topographic domain subdivision (Table 3.4 and Table 3.5), a GWR technique with the Latitude covariate and an additional IDW on residuals resulted in the best fit for San Giacomo Valley Domain, followed by OK, which resulted in the same performance both with an isotropic and anisotropic (direction 135°) variogram of Gaussian type. In the Bregaglia Valley topographic Domain, the model performance was quite poor, with the best performances obtained with OK, with an anisotropic (direction 135°) variogram performing slightly better than the isotropic one (both of Gaussian type).

Table 3.4 Regionalization model parametrization, performance and error metrics results for Jv in the San Giacomo Valley Topographic domain.

Technique	Parametrization	r correlation coefficient observed-simulated	NMRSE
	San Giacomo Valley		
IDW	Idp=2	0.55	0.160
TPS	Covariates=coordinates	0.48	0.164
GWR	Bisquare kernel, adaptive bandwitdth, AIC criterion Covariate=Latitude	0.51	0.162
GWR+IDW on residuals	Bisquare kernel, adaptive bandwitdth, AIC criterion Covariate=Latitude Idp=2	0.57	0.156
OK	Gaussian Isotropic Lag=750 m	0.52	0.159
OK	Gaussian Directional 135° Lag=750 m	0.52	0.159

Table 3.5 Regionalization model parametrization, performance and error metrics results for Jv in the Bregaglia Valley Topographic domain.

Technique	Parametrization	r coefficient	NMRSE
	Bregaglia Valley		
IDW	*Negative correlations for different idp (1.5,2,3,4)		
TPS	Covariates=coordinates+elevations	0.31	0.178
GWR	Bisquare kernel, adaptive bandwitdth, AIC criterion Covariate=Latitude	0.135	0.185
OK	Gaussian Isotropic Lag=500 m	0.35	0.157
OK* ***in the LOOVCV the same variogram (i.e. the general variogram,SILL=0.2, Nuggett=0.05, Range=1000) was applied for each fold, as no convergence in some fold was experienced during the automatic procedure.**	Gaussian Directional 135° Lag=500 m	0.4	0.155

For what concerns the geo-structural domains subdivision, performances resulted always poorer than the Global Domain (Table 3.6 and Table 3.7). The best performing techniques were IDW for the Domain 1 and TPS for Domain 2. The combinations of the best performing maps in the different domains are shown in Figure 3.14. They revealed contrasting values at the borders, up to a difference of 30 in the Jv value. Nevertheless, the spatial pattern was coherent with the one obtained on the Global Domain, especially for the maps in Figure 3.14b and Figure 3.14c, obtained with the OK technique on both the domains.

Table 3.6 Regionalization model parametrization, performance and error metrics results for Jv in the geo-structural Domain 1.

Technique	**Parametrization**	**r coefficient**	**NMRSE**
	Domain 1		
IDW	Idp=2	0.48	0.193
TPS	Covariates=coordinates	0.26	0.200
GWR	Bisquare kernel, adaptive bandwitdth, AIC criterion Covariate=Longitude+elevation	0.24	0.213
GWR+IDW on residuals	Bisquare kernel, adaptive bandwitdth, AIC criterion Covariate=Longitude+elevation Idp=2	0.45	0.200
OK	Circular Isotropic Lag=1300 m	0.42	0.217
OK	Spheric Directional 135° Lag=500 m	0.43	0.216

Table 3.7 Regionalization model parametrization, performance and error metrics results for Jv in the geo-structural Domain 2.

Technique	**Parametrization**	**r coefficient**	**NMRSE**
	Domain 2		
IDW	*correlations always<0.1 for different idp (1.5,2,3,4)		
TPS	Covariates=coordinates+elevations	0.35	0.173
GWR	Bisquare kernel, adaptive bandwitdth, AIC criterion Covariate=Latitude+Longitude+Elevation	0.26	0.179
OK	Gaussian Isotropic Lag=500 m	0.31	0.179

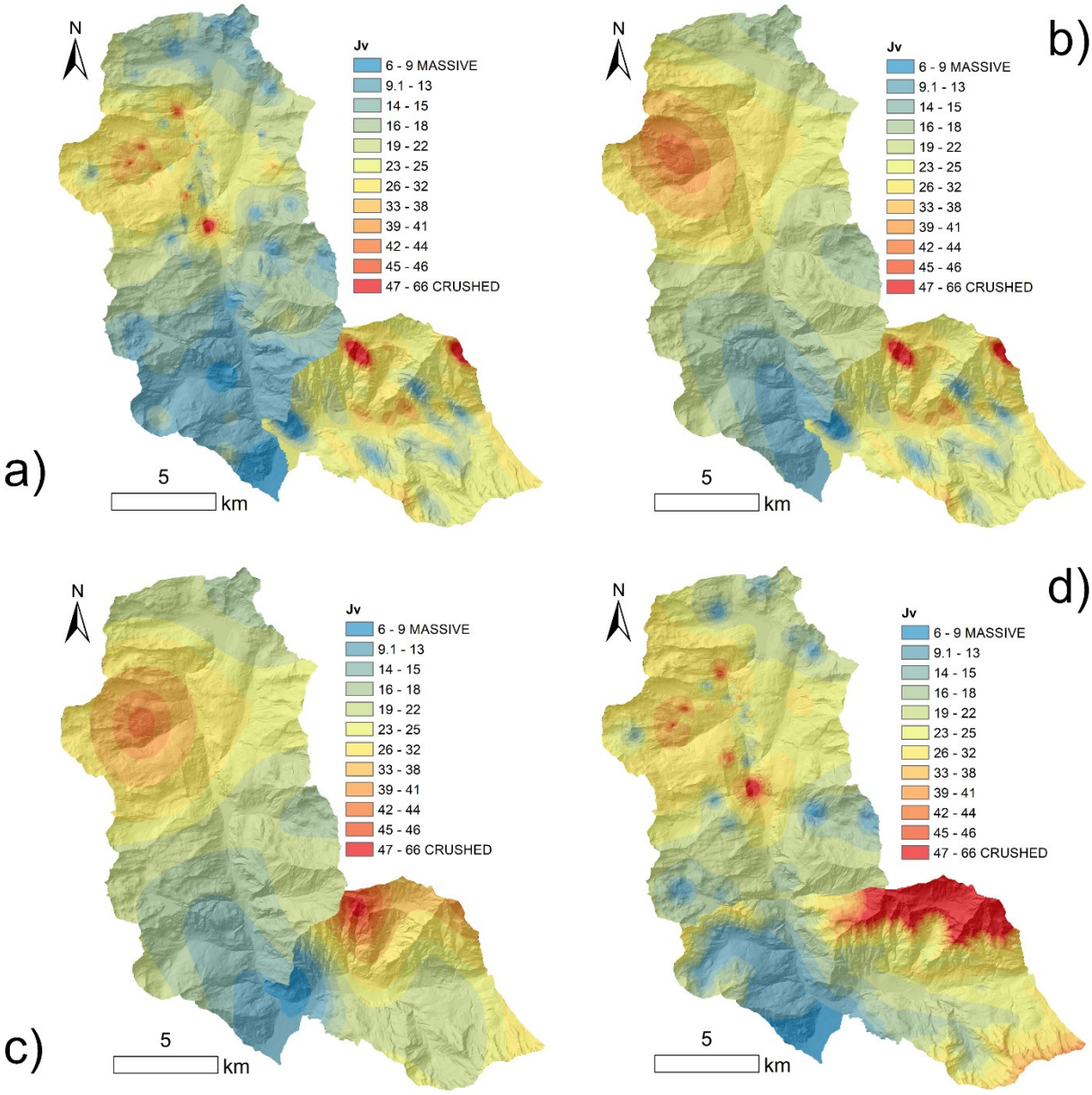


Figure 3.14. Topographic domains: a) Jv map deriving from GWR+IDW in San Giacomo Valley and anisotropic OK in Bregaglia Valley b) Jv map deriving from anisotropic OK in San Giacomo Valley and anisotropic OK in Bregaglia Valley c) Jv map deriving from isotropic OK in San Giacomo Valley and isotropic OK in Bregaglia Valley. Geo-structural domains: d) Jv map deriving from IDW in San Giacomo Valley and TPS in Bregaglia Valley.

For the regionalization of the weathering index in the Global Domain, even if the best performance (Table 3.8) was obtained adopting a GWR technique that included the latitude covariate, the resulting map showed an unrealistically homogeneous spatial distribution of the values (Figure 3.15). For this reason, the map obtained applying an isotropic ordinary kriging (Figure 3.16a-d) was considered more reliable, despite a slightly worse performance (r=0.47, NMRSE=14.5%).

Table 3.8 Regionalization model parametrization, performance and error metrics results for Wi in the Global Domain.

Technique	Parametrization	r coefficient	NMRSE
IDW	Idp=2	0.44	0.147
TPS	Covariate=coordinates	0.44	0.152
GWR	Bisquare kernel, adaptive bandwitdth, AIC criterion Covariate=Latitude	0.51	0.140
OK	Exponential Isotropic Lag=750 m	0.47	0.145

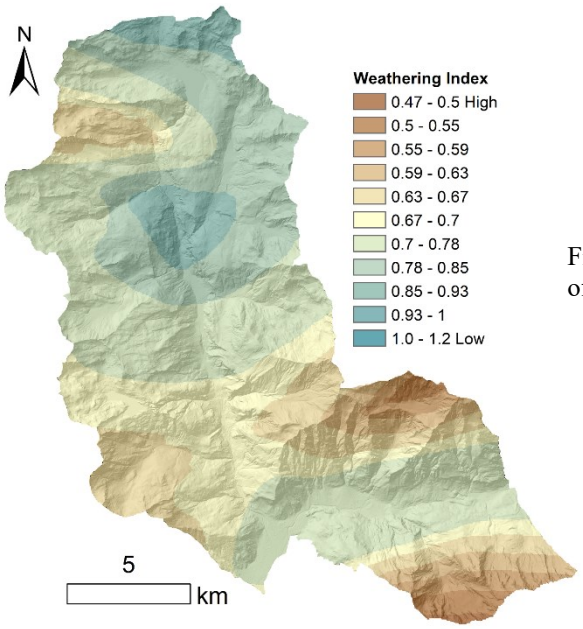


Figure 3.15 Wi regionalizartion deriving from the application of the GWR technique with the latitude covariate.

The variography of Wi (Figure 3.16c) resulted in an isotropic spherical variogram with a nugget effect of 0.0075 and a partial sill of 0.008. It is noteworthy that the nugget and the partial sill were very similar, thus being Wi at the edge of the definition of regionalized variable. The shorter range (i.e., 3125 m) and lag (i.e., 750 m) in comparison to Jv confirmed the expected more local connotation of this property in relation to the fracturing grade, as Wi may express the local microclimate, while Jv is the reflection, at the outcrop scale, of the regional structural and tectonic framework. Sill and range values remained consistent amongst the LOOV-CV sets (Figure 3.16d). These results are geologically plausible, as no evidence of a weathering grade anisotropy was ever observed or reported in field surveys.

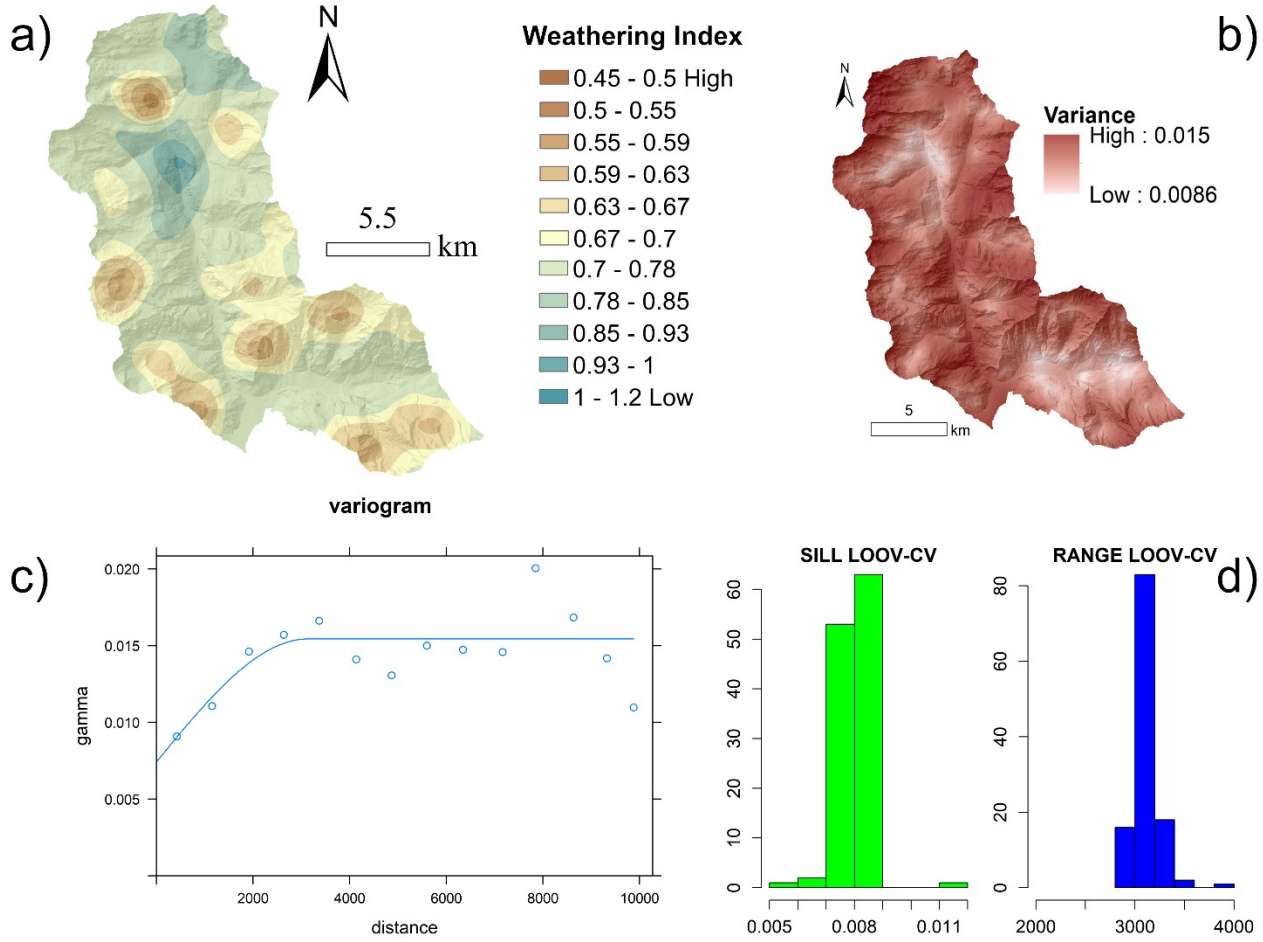


Figure 3.16. a) Wi map deriving from the isotropic ordinary kriging applied on the Global Domain and b) associated variance map. c) experimental variograms of Wi (points) and variogram models (line) d) histograms representing the variations of Sill and Range coming from the different LOOV-CV sets.

Regarding the topographic domain subdivision, for the San Giacomo Valley Domain a KED technique with the Latitude covariate performed best, followed by a GWR technique with the Latitude covariate and an additional IDW on the residuals (Table 3.9). In the Bregaglia Valley topographic Domain (Table 3.10), a KED technique with the Elevation covariate performed best, followed by a GWR technique with the Elevation covariate.

Table 3.9 Regionalization model parametrization, performance and error metrics results for Wi in the San Giacomo Valley Topographic domain.

Technique	Parametrization	r coefficient	NMRSE
	San Giacomo Valley		
IDW	Idp=2	0.50	0.175
TPS	Covariates=coordinates	0.42	0.189
GWR	Bisquare kernel, adaptive bandwitdth, AIC criterion Covariate=Latitude	0.47	0.178
GWR+IDW on residuals	Bisquare kernel, adaptive bandwitdth, AIC criterion Covariate=Latitude Idp=2	0.51	0.176
OK	Exponential Isotropic Lag=500 m	0.47	0.179
KED	Exponential Isotropic Lag=500 m Covariate=Latitude	0.52	0.173

Table 3.10 Regionalization model parametrization, performance and error metrics results for Wi in the Bregaglia Valley Topographic domain.

Technique	Parametrization	r correlation coefficient observed-simulated	NMRSE
	Bregaglia Valley		
IDW	Idp=5	0.24	0.199
TPS	Covariates=coordinates+elevations	0.45	0.167
GWR	Gaussian kernel, adaptive bandwitdth, AIC criterion Covariate=Elevation	0.51	0.156
OK	No variogram model converged. Evident drift in the experimental variogram (i.e. linear trend)		
KED	Exponential Isotropic Lag=300 m Covariate=elevation	0.56	0.151

For what concerns the geo-structural domains subdivision, in Domain 1 a GWR technique with the Latitude covariate and an additional IDW on residuals performed best, followed by a IDW technique (Table 3.11). In Domain 2, a KED technique with the Elevation covariate performed best, followed by a GWR technique with the Elevation covariate (Table 3.12).

Table 3.11 Regionalization model parametrization, performance and error metrics results for Wi in the geo-structural Domain 1.

Technique	**Parametrization**	**r coefficient**	**NMRSE**
	Domain 1		
IDW	Idp=2.5	0.55	0.176
TPS	Covariates=coordinates	0.51	0.180
GWR	Gaussian kernel, adaptive bandwitdth, AIC criterion Covariate=Latitude	0.45	0.200
GWR+IDW on residuals	Gaussian kernel, adaptive bandwitdth, AIC criterion Covariate=Latitude Idp=2	0.57	0.174
OK	Circular Isotropic Lag=600 m	0.50	0.181

Table 3.12 Regionalization model parametrization, performance and error metrics results for Wi in the geo-structural Domain 2.

Technique	**Parametrization**	**r coefficient**	**NMRSE**
	Domain 2		
IDW	*correlations always<0.1 for different idp (1.5,2,3,4)		
TPS	Covariates=coordinates+elevation	0.35	0.185
GWR	Gaussian kernel, adaptive bandwitdth, CV criterion Covariate=Elevation	0.46	0.166
OK	Gaussian Isotropic Lag=500 m	0.25	0.188
KED	Exponential Isotropic Lag=300 m Covariate=elevation	0.47	0.166

The combinations between the best performing maps consisted of four maps for the Topographic subdivision (Figure 3.17) and four maps for the geo-structural subdivision (Figure 3.18).

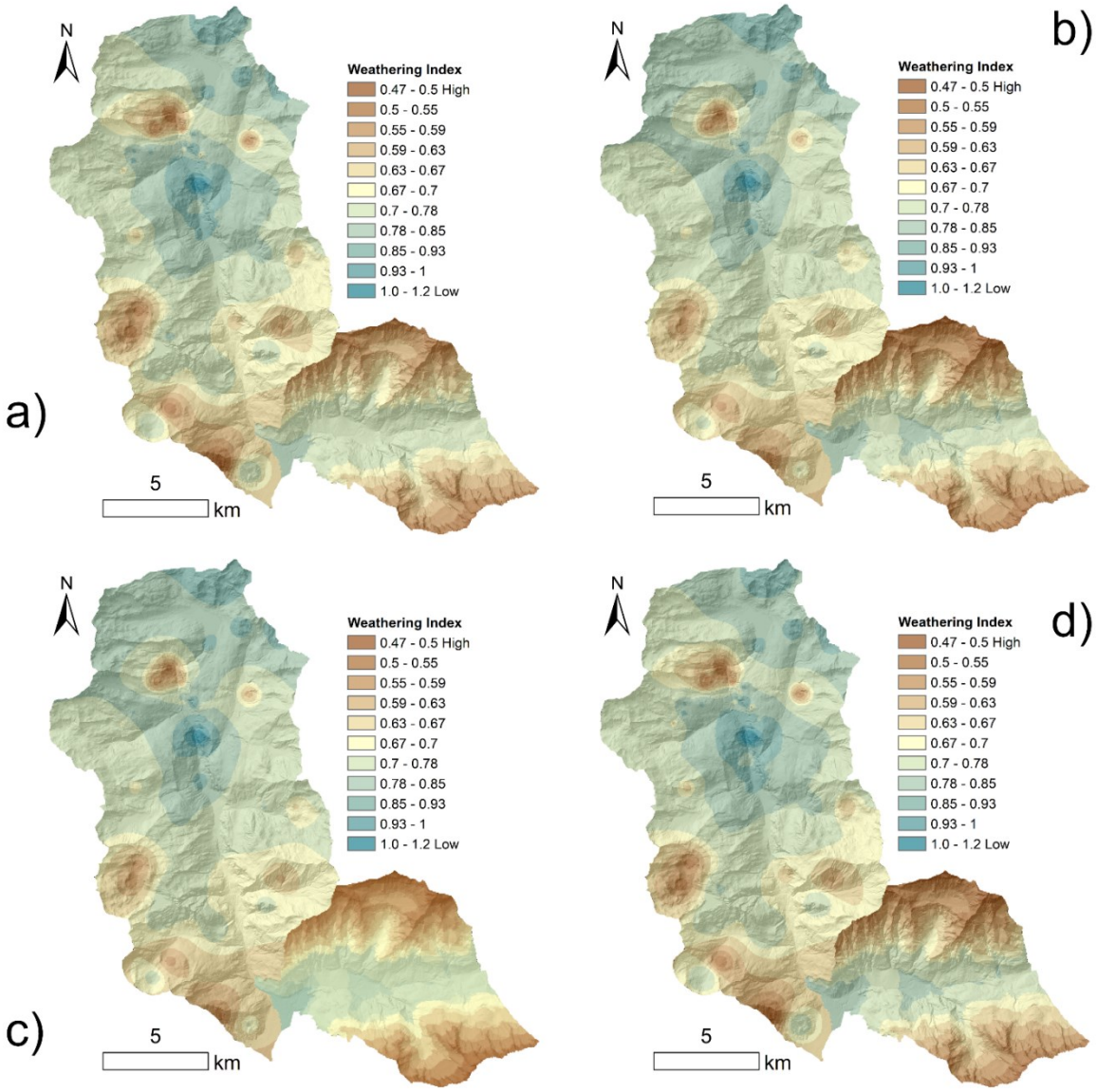


Figure 3.17. Topographic domains: a) Wi map deriving from the application of GWR in San Giacomo Valley and GWR in Bregaglia Valley b) Wi map deriving from the application of KED in San Giacomo Valley and KED in Bregaglia Valley c) Wi map deriving from the application of GWR in San Giacomo Valley and KED in Bregaglia Valley. d) Wi map deriving from the application of KED in San Giacomo Valley and GWR in Bregaglia Valley.

The spatial patterns of the combined Wi maps revealed less visible, although still present, incoherencies at the domain borders than Jv maps. In general, all the maps resulted very similar both mutually and if compared to the Global Domain maps. Independently from the regionalization domain, the weathering index, especially in Bregaglia Valley, showed a clear increasing trend with elevation, which was interpreted as the result of a combination of exposure to freeze-thaw cycles, scarce vegetation, and glacial history. During the LGM, the glacier did not reach the slopes at the highest elevations, which remained exposed to climate agents.

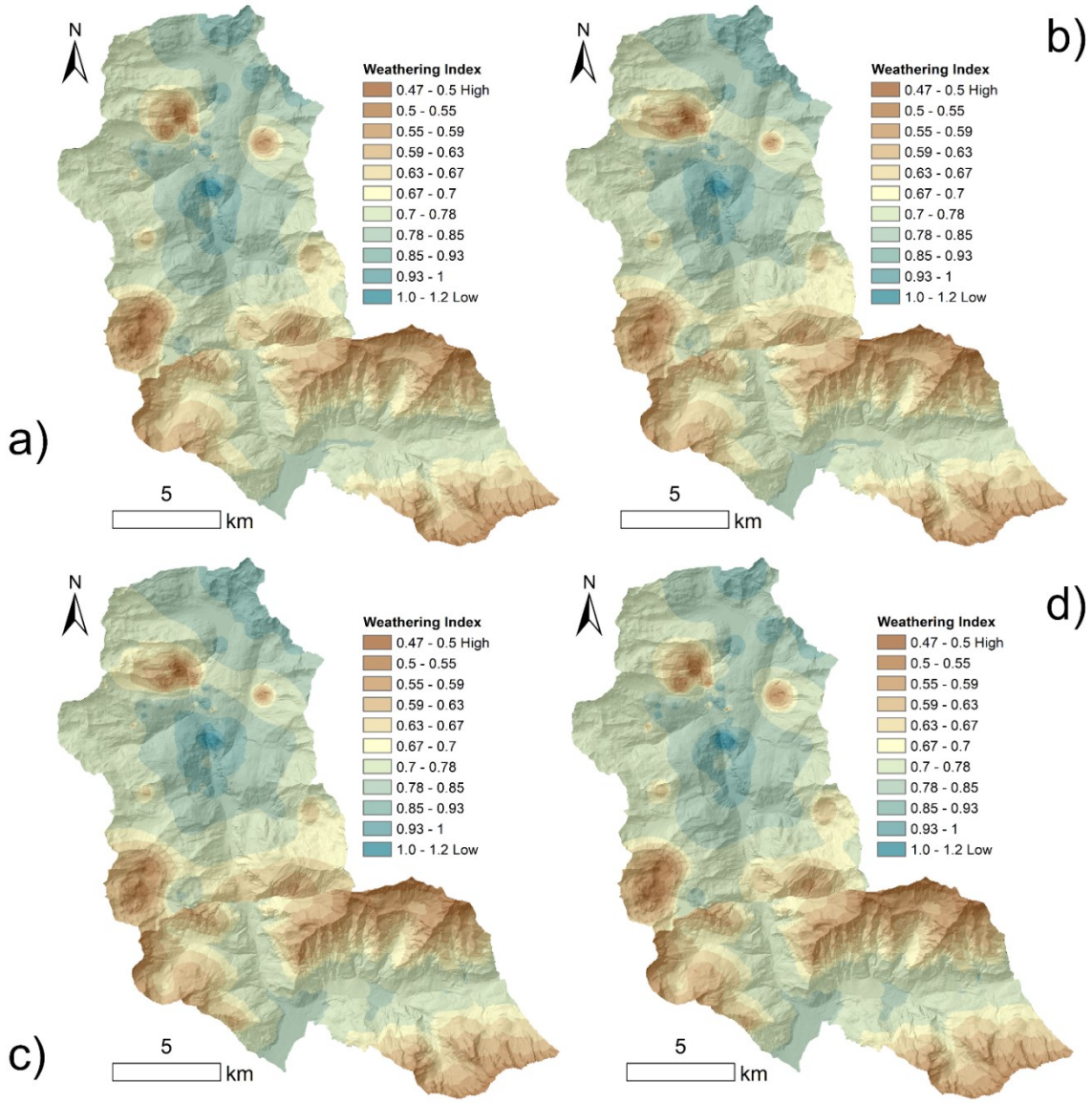


Figure 3.18. Geo-structural domains: a) Wi map deriving from the application of IDW in San Giacomo Valley and GWR in Bregaglia Valley b) Wi map deriving from the application of GWR in San Giacomo Valley and GWR in Bregaglia Valley c) Wi map deriving from the application of GWR in San Giacomo Valley and KED in Bregaglia Valley. d) Wi map deriving from the application of IDW in San Giacomo Valley and KED in Bregaglia Valley.

The regionalization of the Keq index showed the poorest performance and the highest uncertainty among the three properties analysed in this study. This result was consistent considering both the global and the split domains. This may be attributed to the joints' aperture, which is included in the Keq tensor calculation. Variability in aperture is one of the most challenging problem in slope design and management (Price, 2016) and it cannot be considered neither constant nor scale invariant (Baghbanan and Jing, 2008). Aperture is known to vary on short distances, even among the same joint family and along the same joint, especially because of roughness effects along joint surfaces. Its best output on the Global Domain (Table 3.13 and Figure 3.19) was obtained through a TPS method including altitude as covariate (r=0.16, NMRSE=18.5%).

Table 3.13. Regionalization model parametrization, performance and error metrics results for Keq in the Global Domain.

Technique	Parametrization	r coefficient	NMRSE
IDW	Idp=2	0.14	0.178
TPS	Covariate=coordinates+elevation	0.16	0.185
GWR	Gaussian kernel, fixed bandwitdth, AIC criterion Covariate=Longitude+Elevation	0.09	0.180
GWR+IDW on residuals	Gaussian kernel, fixed bandwitdth, AIC criterion Covariate=Longitude+Elevation Idp=2	0.15	0.183
OK	Exponential Isotropic Lag=600 m	0.09	0.180

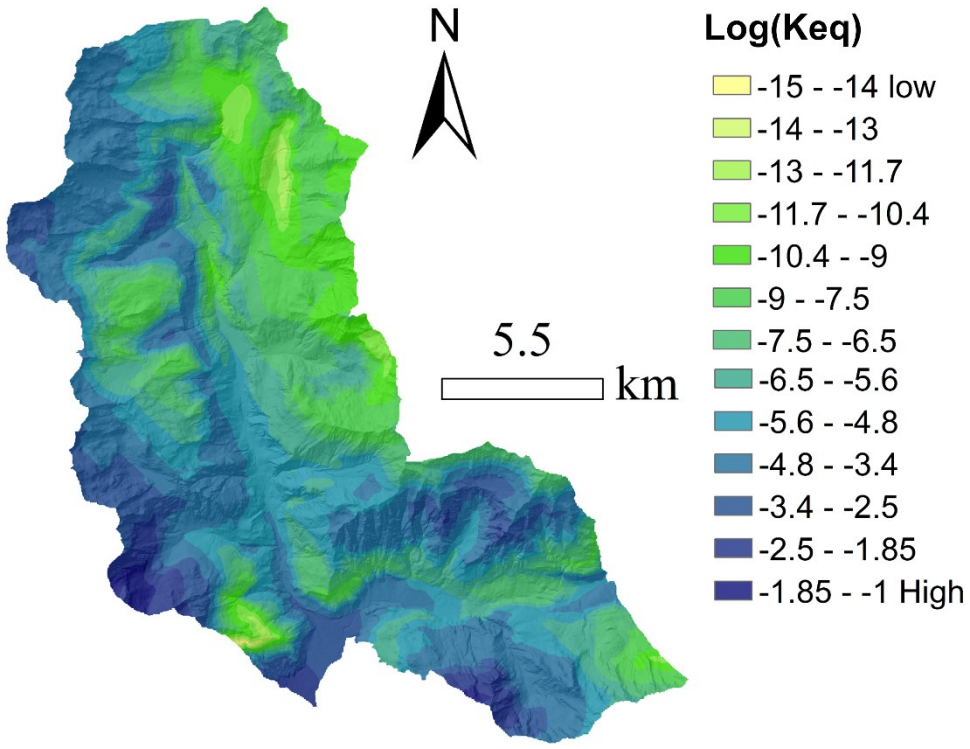


Figure 3.19 Keq regionalizartion deriving from the application of the TPS technique with the altitude covariate.

Regarding the topographic subdivision, for San Giacomo Valley, the best performance, although very poor, was obtained using a GWR with Latitude (Table 3.14), but the spatial pattern was physically implausible, with an almost vertical gradient from W to E and a strong mismatch at the border with the Bregaglia Valley domain (Figure 3.20a). Conversely, for Bregaglia Valley Domain (Table 3.15), both TPS with altitude as a covariate and IDW yielded maps with a correlation coefficient (r) larger than 0.5. Over the Bregaglia Valley, the spatial pattern was very similar to that obtained working with the Global

Domain and TPS (Figure 3.20a). Thus, the poor performance obtained on the Global domain could be mostly imputed to the San Giacomo Valley Domain related poor regionalization performance.

Table 3.14. Regionalization model parametrization, performance and error metrics results for Keq in the San Giacomo Valley Topographic domain.

Technique	Parametrization	r coefficient	NMRSE
	San Giacomo Valley		
IDW	*correlations always negative for different idp (1.5,2,3,4)		
TPS	Covariates=coordinates+slope	0.17	0.198
GWR	Gaussian kernel, adaptive bandwitdth, CV criterion Covariate=Longitude	0.23	0.190
GWR+IDW on residuals	Gaussian kernel, adaptive bandwitdth, CV criterion Covariate=Longitude Idp=1.5	0.13	0.203
OK	*correlations always negative for different variogram models and lags.		

Table 3.15. Regionalization model parametrization, performance and error metrics results for Keq in the Bregaglia Valley Topographic domain.

Technique	Parametrization	r coefficient	NMRSE
	Bregaglia Valley		
IDW	Idp=5	0.55	0.200
TPS	Covariates=coordinates+elevations	0.53	0.190
GWR	Bisquare kernel, adaptive bandwitdth, AIC criterion Covariate=Latitude+Aspect	0.30	0.218
GWR+IDW on residuals	Bisquare kernel, adaptive bandwitdth, AIC criterion Covariate=Latitude+Aspect Idp=6	0.44	0.252
OK	No variogram model converged.		

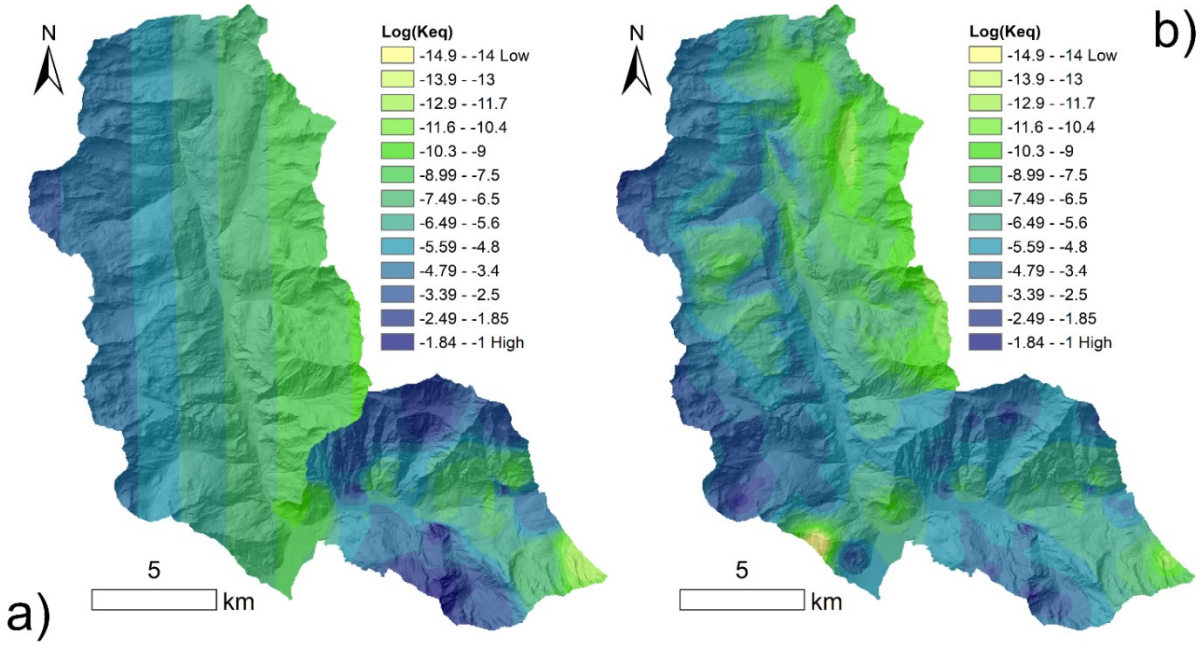


Figure 3.20. Topographic domains: a) Keq map deriving map deriving from the application of GWR in San Giacomo Valley and IDW in Bregaglia Valley. Geo-structural domains: b) Keq map deriving map deriving from the application of TPS in Domain 1 and IDW in Domain 2.

Regarding the geo-structural subdivision, the only significant result was obtained combining TPS with Elevation as covariate for Domain 1 (Table 3.16) and IDW for Domain 2 (Table 3.17). The spatial pattern for both domains approximated well the Global Domain (Figure 3.20b). Furthermore, the performance resulted similar (slightly better for Domain 2 while slightly scarcer for Domain 1), if compared to the performance of the Global Domain regionalization.

Table 3.16. Regionalization model parametrization, performance and error metrics results for Keq in the geo-structural Domain 1.

Technique	Parametrization	r coeffient	NMRSE
	Domain 1		
IDW	*correlations always <0.1 for different idp (1.5,2,3,4)		
TPS	Covariates=coordinates+elevation	0.13	0.230
GWR	*correlations always negative for different covariates and model parameters combinations		
OK	*correlations always negative for different variogram models and lags combinations		

Table 3.17 Regionalization model parametrization, performance and error metrics results for Keq in the geo-structural Domain 2.

Technique	Parametrization	r coefficient	NMRSE
	Domain 2		
IDW	Idp=2	0.24	0.18
TPS	Covariates=coordinates+elevation	0.153	0.192
GWR	*correlations always negative for different covariates and model parameters combinations		
OK	Gaussian Isotropic Lag=500 m	0.11	0.183

In summary, the regionalized geomechanical properties, chosen as predictors for the consequent rockfall susceptibility model, were the ones related to the Global Domain. For Jv, the map associated with the anisotropic ordinary kriging was selected, while for Wi the one associated with the isotropic ordinary kriging. For Keq, the map associated with a TPS with the altitude covariate was chosen.

3.5.2. Rockfall susceptibility modelling with Generalized Additive Models

3.5.2.1 Predictors response relationships

Variable penalization resulted in the exclusion of the predictors eastness and plan curvature from the TOPO model fitted on all data and in most sCV runs (selection frequencies 33.0% and 13.4%, respectively). Therefore, these predictors were excluded from the successive GM and GM+GEO models. A focus on the GM+GEO model results is presented in the current section, while the observed differences between the two models would be discussed at the end of the Section. For the discussion below, refer to Figure 3.21 for predictors component smoothing functions (CSF), and Figure 3.22 for *edf* (used for the calculation of the penalization frequency, see Section 3.4.5.1) and the mDD% variations amongst the different CV runs.

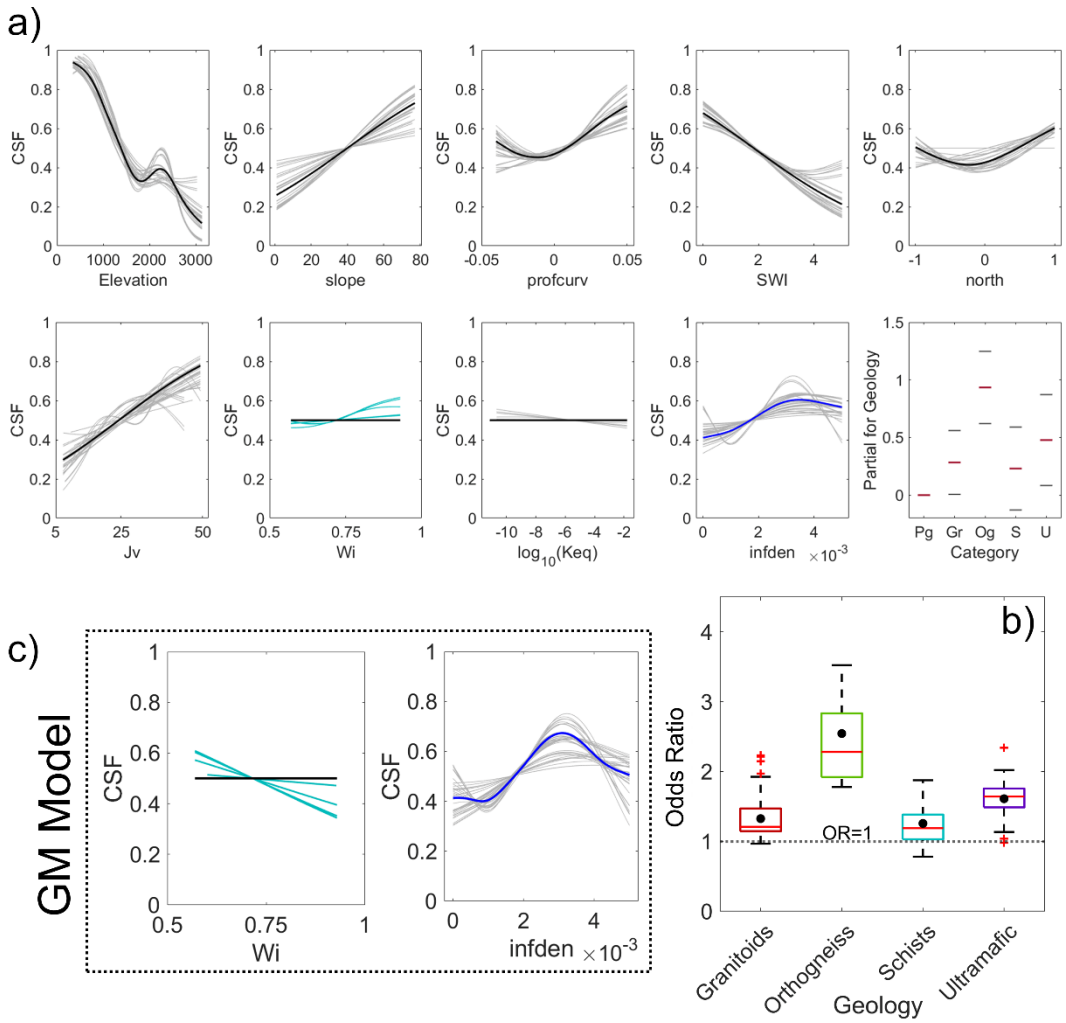


Figure 3.21Smoothing functions and odds ratios extracted from all the sCV runs (grey lines –light blue for Wi – for smoothing functions and boxplot for odds ratios) and from the model fitted on all data (black line – blue for infden – for smoothing functions and black dots for odds ratios). a) Smoothing function for the predictors included in the GM+GEO Model b) odds ratios for the geological predictor c) Wi and Infiltration density smoothing functions from all the sCV runs and from the model fitted on all data for the GM Model

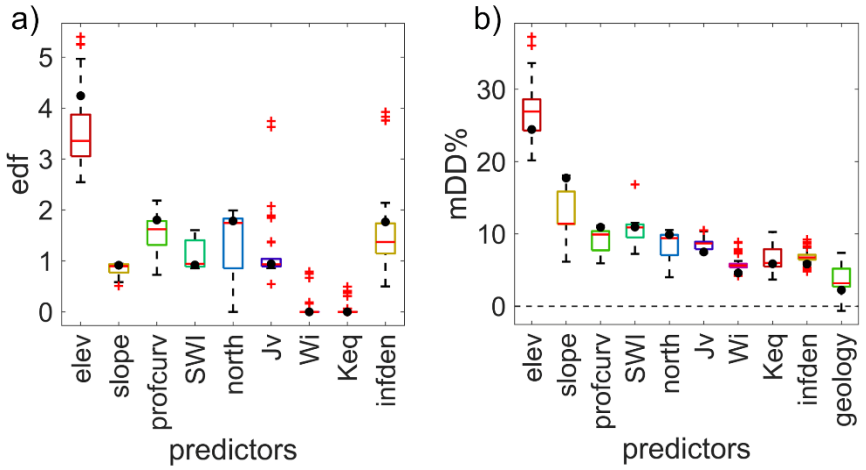


Figure 3.22 a) Effective degrees of freedom (edf) and b) mDD% for each predictor for the GM+GEO model. Boxplots represent values from the sCV runs, while superimposed black dots are estimated using all data. Variables abbreviations: elev=elevation; profcurv=profile curvature; north=northness; infden=infiltration density.

The predictor-response relationships (Figure 3.21a) of the morphometric predictors, Jv and Keq were maintained through the GM and the GM+GEO models, while some variations in Wi and Infiltration Density behaviour between the two models were observed (Figure 3.21c).

In the GM+GEO model, elevation and slope were the most important predictors with high selection frequencies of 100% and 80.4%, and mDD% values of 24.5% and 17.7% on the entire dataset. Rockfall susceptibility decreased quite strongly towards higher elevations consistently with the valley morphology, which is highly incised, with active rock walls at the bottom (Figure 3.23). Indeed, in deglaciated valley, glacial erosion may steepen rock slopes, increasing overburden and deepening scarp-foot locations (Ballantyne, 2002); on the other hand, deglaciation induced compression along the valley bottom increasing differential stresses and causing rock damage.

The secondary small increase of susceptibility between 2000 and 2200 m a.s.l. could be linked to the transition between glacial terraces, characterized by soil-dominated slopes, and high elevation sub-vertical rock walls. Coherently, rockfall susceptibility increased towards higher slope angles. Profile curvature, SWI and northness were almost always included (100%, 100%, and 80.4%, respectively) and showed a mDD% in the all-points based model around and 11% for both profile curvature and SWI and 10% for northness. Regarding geomechanical predictors, rockfall susceptibility increased linearly with increasing Jv, both in the all-points-based model and in the sCV runs (sCV selection frequency 96.4%, mDD% 7.5% using all data). Rock mass fracturing degree therefore had a quite strong influence on rockfall susceptibility predictions, comparable to topographic predictors other than elevation and slope. The modelled functional relationship remained stable across sCV runs, indicating that the relationship is spatially transferable within the study area.

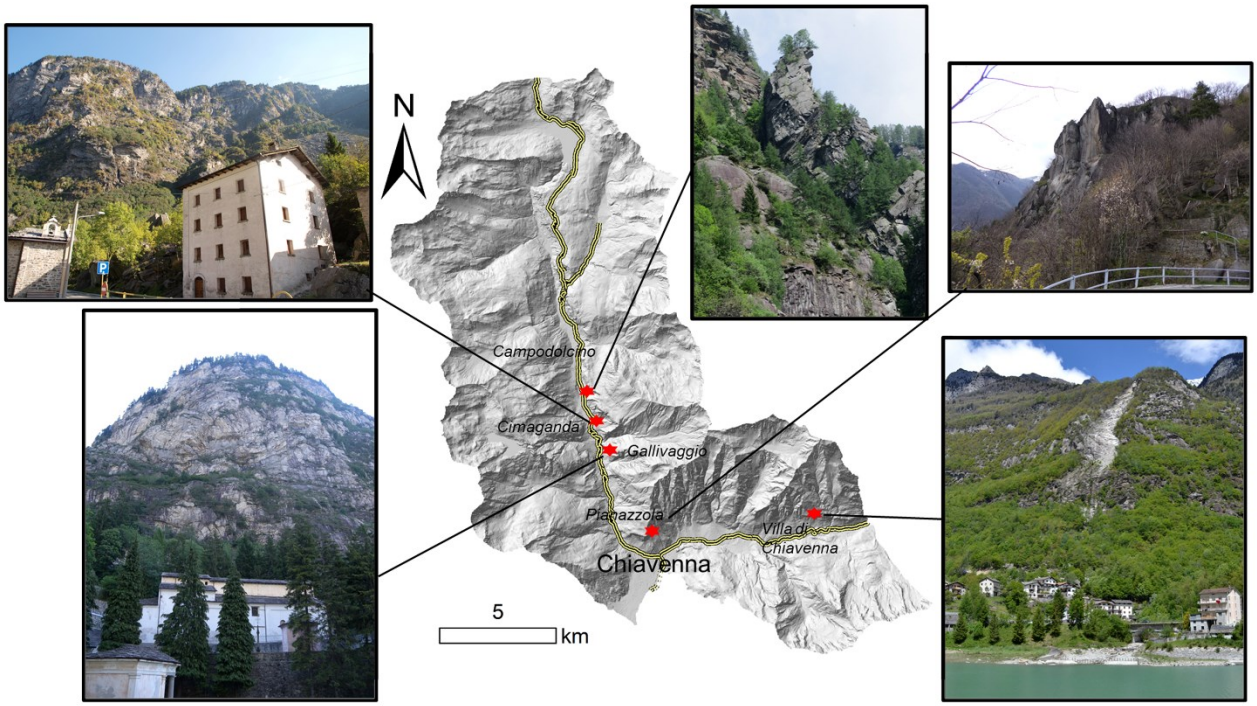


Figure 3.23 Photographs and location of some active rockwalls at the valley bottoms.

Wi and Keq demonstrated a more unstable behaviour. They were penalized out of the model using all data (edf equal to 0), and in the 93.4% and 100% of the sCV runs respectively (respect to the fixed threshold of edf<0.7). The exceptions are represented by the outliers with edf > 0 in Figure 3.22a.

The mDD% consistently showed values below 10%, although never approaching zero. Overall, these two predictors can be considered of minor importance. Although Wi regionalization had a good performance, its values showed a limited range (from 0.5 to 1.0), which may indicate that the area was not subject to a sufficient differential weathering to influence globally rockfall occurrence. On the other hand, the influence of the proximity to Last Glacial Maximum trimlines (refer to Figure 3.3 in Section 3.2), which usually causes high weathering (Matasci et al., 2011), may be masked by the presence of the elevation predictor. Its secondary peak is indeed compatible with the trimline location. Nonetheless, the trend of the predictor smoothing function in the few sCV runs where Wi was not penalized was not physically reliable, as it showed higher susceptibility for less weathered areas (note that Wi→1 means low weathering). For Keq, the model was not able to capture the multifaceted aspect of the processes associated to the property (see Section 3.4.4 for details). Also, Keq may be more suitable for local-scale analyses rather than for regional scale generalizations, especially because its calculation is based on the hydraulic aperture of the joints, which could be highly variable over short distances, due to the presence of asperities at different scales modulating the effective volume available for water flow. However, it

was still possible to consider the hydrogeological influence on rockfall susceptibility using the infiltration density predictor. This variable was included in almost all sCV runs (selection frequency 94.8%) and achieved an mDD% of 5.8% using all data (Figure 3.22a and Figure 3.22b). The associated smoothing function showed an increasing behaviour for low to medium values and the reaching of a plateau at a value of approximately $3x10^{-3}$ number of elements per m^2 both for the all-points-based model and for most of the sCV runs. In few cases the curve resulted in a bell shape, not physically meaningful and possibly attributable to the train-test splitting (Figure 3.21a).

Considering the geological predictor, with an odds ratio of 2.5 in all-points-based model, orthogneiss lithology was associated with a statistically significative (*p*-value of 0.002) higher chance of rockfall occurrence than the most abundant category (i.e., paragneiss). All the other lithologies showed odds ratios close to 1 meaning a similar chance of rockfall occurrence as the paragneiss lithology (Figure 3.21b). Consequently, the overall mDD% of geology was relatively low (2.3% with all data, up to 7.4% in sCV runs). One may argue that orthogneiss are in general a more resistant lithology than paragneiss. However, in the study area, orthogneisses mainly belong to the Gruf Complex, which has peculiar petrographic and structural characteristics. It is a poly-metamorphic highly foliated migmatitic body, with different facies and deformation structures both at the meso- and micro-scale (Galli et al., 2013). This led to inherent mechanical and mineralogical contrasts, as well as weak zones, probably not fully captured by the Jv predictor. As demonstrated by the recent work by De Vilder et al. (2017), not only the wider geology but also the local scale lithological and mineralogical variability controls rock mass wasting processes, as observed in the Gruf Complex. Moreover, the presence of the Gruf Line in this area represents a pervasive weak zone particularly prone to instability.

Within the GM+GEO model, the concurvity among continuous predictors resulted to be acceptable, with a maximum value of 0.59. Although almost identical to the GM+GEO model discussed above, the GM model was characterized by two differences in terms of predictors behaviour (Figure 3.21c). Firstly, in some sCV runs the Wi predictor had an opposite behaviour in comparison to the GM+GEO model and in this case the curve trend was physically coherent with the process (i.e., higher susceptibility for more weathered areas). Secondly, the infiltration density predictor was characterized by the prevalence of a bell shape, even if the plateau-type shape observed in the GM+GEO model was preserved in some sCV runs. This may be attributable to a large variability of the dataset at the highest values, potentially smoothed with the introduction of geology, rather than to a real physical meaning. Furthermore, all the geomechanical predictors and infiltration density had a higher deviance explained than in the GM+GEO model (of the order of 2-3% higher).

3.5.2.2 Model performance

Figure 3.24 shows the performance of the three models. The prediction performance of the TOPO model (mean AUROC of 0.68 in sCV) indicated a poor discrimination capacity, while those of the GM and GM+GEO models (mean AUROC of 0.71 and 0.72, respectively) can be regarded as acceptable, according to Hosmer et al. (2013) guidelines. Considering the IQR of AUROC across sCV runs, the GM model showed the best spatial transferability, followed by GM+GEO. This may be linked to the fact that the geology in the study area is spatially compartmentalized, leading to a higher heterogeneity between training and test samples and thus to a lower transferability (Guzzetti et al., 2006; Petschko et al., 2014). Besides the governing role of gravity in rockfall occurrence, these results highlighted that geomechanical (Jv) and geomorphological-hydrogeological (infiltration density) predictors improved model performance, allowing a deeper and more process-oriented understanding of the phenomena, crucial for the practical use of susceptibility maps.

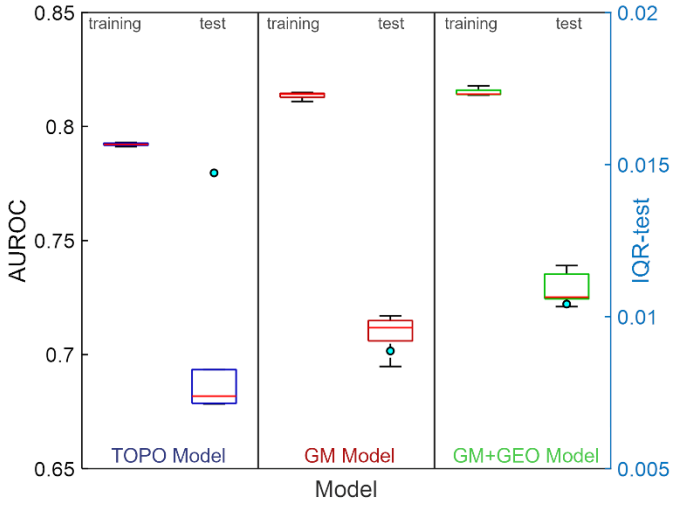


Figure 3.24 Performance of the three models (TOPO, GM and GM+GEO) in the sCV runs (repetitions), expressed in terms of AUROC values (boxplots, primary y-axis) for both the training and the test set and IQR (points, secondary y-axis) for the test set only.

3.5.2.3 Susceptibility map and geomorphological plausibility

The three output susceptibility maps are shown in Figure 3.25. The susceptibility spatial pattern in the TOPO model showed a concentration of the "very high" and "high" susceptibility classes along the cliffs located at the base of the principal valley bottoms. The percentages of the "very high" and "high" classes in the GM model remained almost equal to the TOPO model. However, their spatial distribution changed, including also some lateral valleys and rock walls and ridges at higher elevations. The spatial pattern in the GM+GEO model was very similar to the previous GM model, except for a slightly higher susceptibility in correspondence of the Gruf Complex in the southern slope of the Bregaglia Valley. Both the GM and GM+GEO models showed a better discrimination than the TOPO model, as the "medium"

susceptibility class, which can be considered as the most ambiguous, was less widespread in both models.

By qualitatively comparing the GM model output map with the rock walls defined as “active” and the rock masses defined as “very shattered” in the available geomorphological maps of the area, in most cases a good agreement with the “very high” and “high” susceptibility classes was observed Figure 3.25d-i). Some mismatches were still present, probably due to the different scales of the survey (1:10,000) and the geological map used to discriminate rock masses and quaternary deposits (1:50,000). In some areas, limitations related to restricted visibility (e.g., complex topography, dense vegetation) may also have affected field mapping.

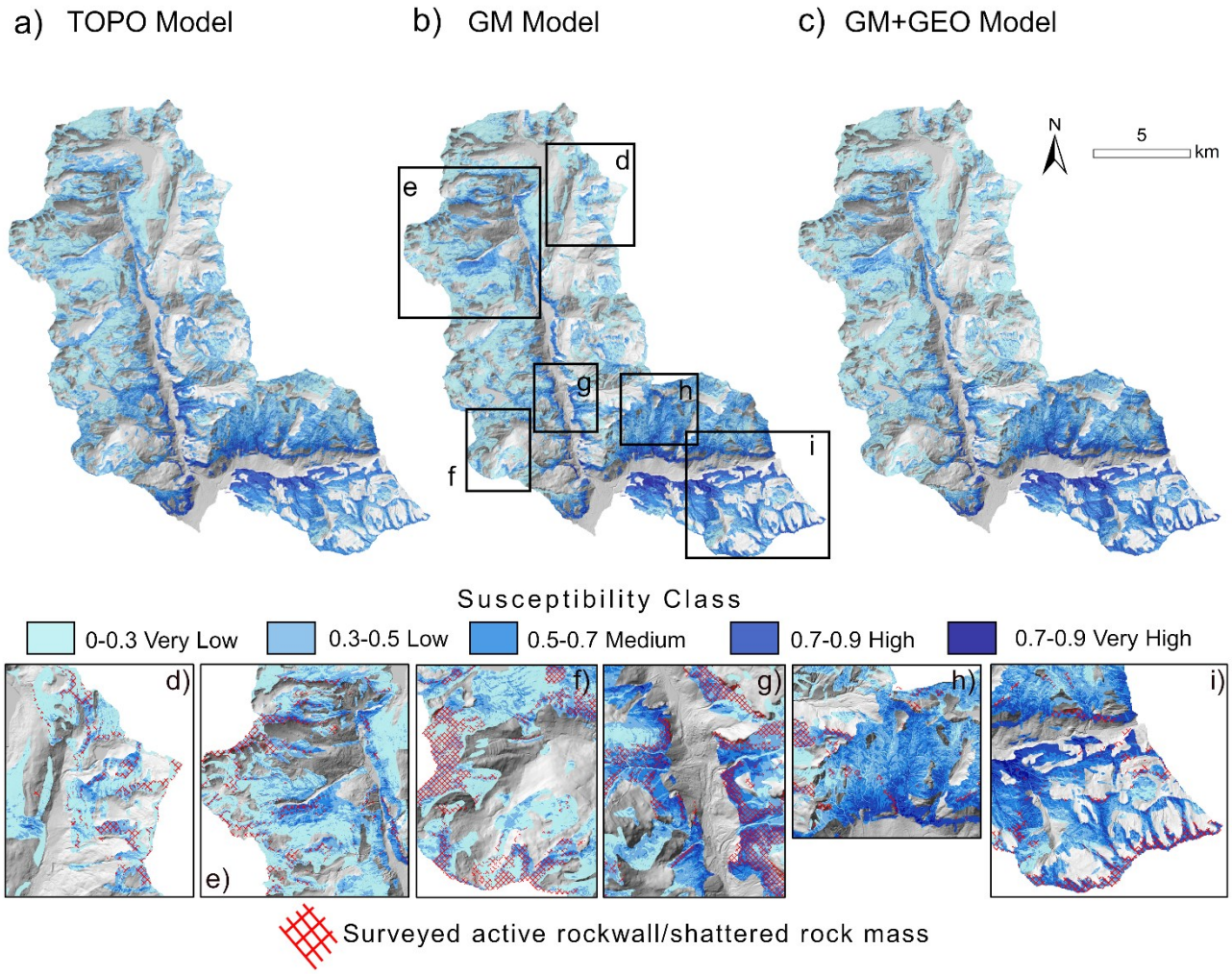


Figure 3.25 Susceptibility maps obtained from a) topographic (TOPO), b) geomechanical (GM), and c) geomechanical-geological (GM+GEO) models; d)-i) zoom-in of the geomechanical (GM) susceptibility map with superimposed active rock walls surveyed in the field.

Figure 3.26a shows the differences between the maps obtained with the TOPO and GM models (GM-TOPO). This map pointed out how the introduction of the geomechanical predictors led to an increased susceptibility in some geomorphologically peculiar areas. The first of these areas is in proximity to the DSGSD of the Febbraro and Vamlera secondary valleys, located in the upper western part of San Giacomo Valley (Figure 3.26c-d-e). The second area is located along the ridge drawn by Mt. Emet (Figure 3.26f) and Mt. Mater, located in the upper Eastern part of the San Giacomo Valley. This area is well-known and monitored from an engineering geology point of view, since it is characterized by two interacting DSGSDs, associated to secondary rockslides and rock glaciers (Crippa et al.,2020). The last area crosses the regional thrust separating the Suretta and Tambò nappes on the northern slope of the Bregaglia Valley (Figure 3.26g). The decrease in susceptibility in the central-lower part of San Giacomo Valley is linked either to a combination of low fracturing grades and to very low infiltration density values or to the low confidence of the model for very high infiltration density values (see Section 3.5.2.1). In particular, the decrease in susceptibility in the SW part of the study area may be referred to low values of Jv (i.e., more massive rock masses) modelled (and observed) in this part of the territory.

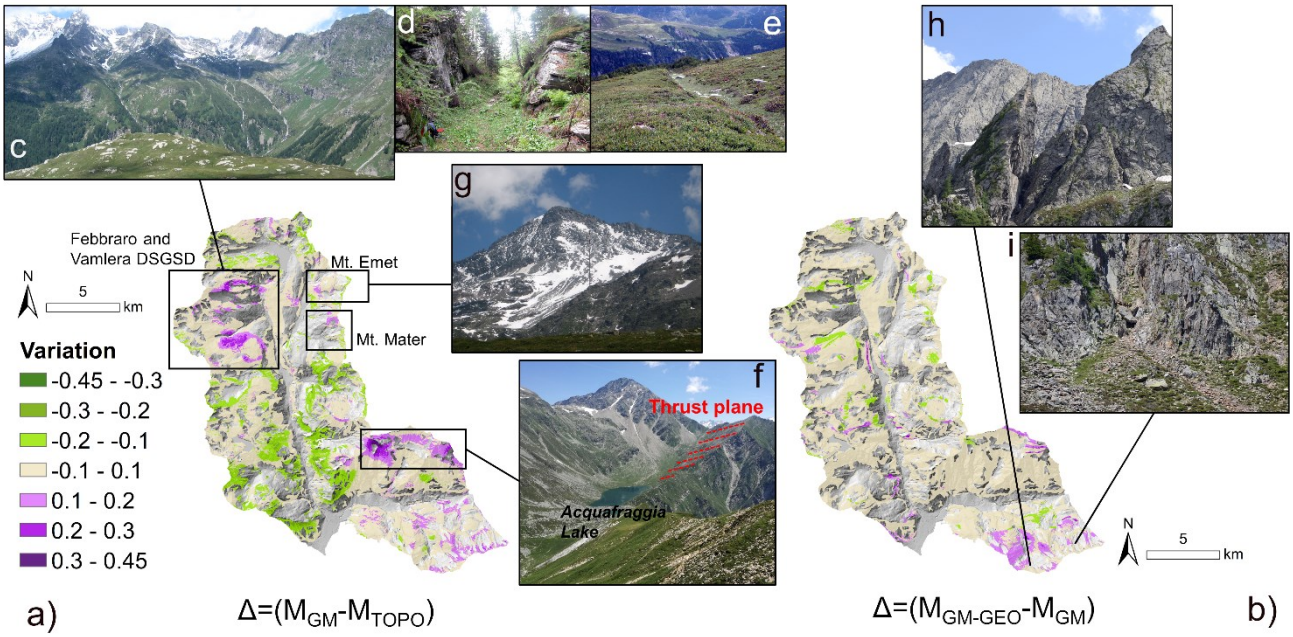


Figure 3.26. a) Difference map between GM model and TOPO model b) Difference map between GM+GEO model and GM model c),d),e),f),g),h),i) photograph of the mentioned locations, courtesy of the Geoengineering group of University of Milan.

From a geological-geomechanical point of view, this area corresponds to a pervasive and continuous presence of the Truzzo Granite, which has a variable structure from mesoscopic lenses of undeformed granite to highly strained orthogneiss (Marquer et al. 1994; Carlà et al., 2019). The occurrence of low-medium values of Jv are probably associated to the prevalence of this second facies. However, as highlighted in Figure 3.25, this does not mean that these slopes are totally not susceptible, with portions

still in the "medium" and "high" classes, rather that, in relative terms, the geomechanical component of the model favours stability. Figure 3.26b shows the variation map between the GM and the GM+GEO models. This map highlighted a further increase of susceptibility in the southern part of the Bregaglia Valley, due to the peculiar characteristics (mylonitic rocks, ductile shear zones, mechanical contrasts) of the Gruf Complex (Figure 3.26 h-i), already discussed in Section 3.2.

The comparison between the difference map between the TOPO and GM models with the PS/DS mean annual velocities, obtained from SqueeSAR™ data for the period 2014-2021 (Figure 3.27a-c), showed that all the above-mentioned areas are associated with recent deformation and displacements along the

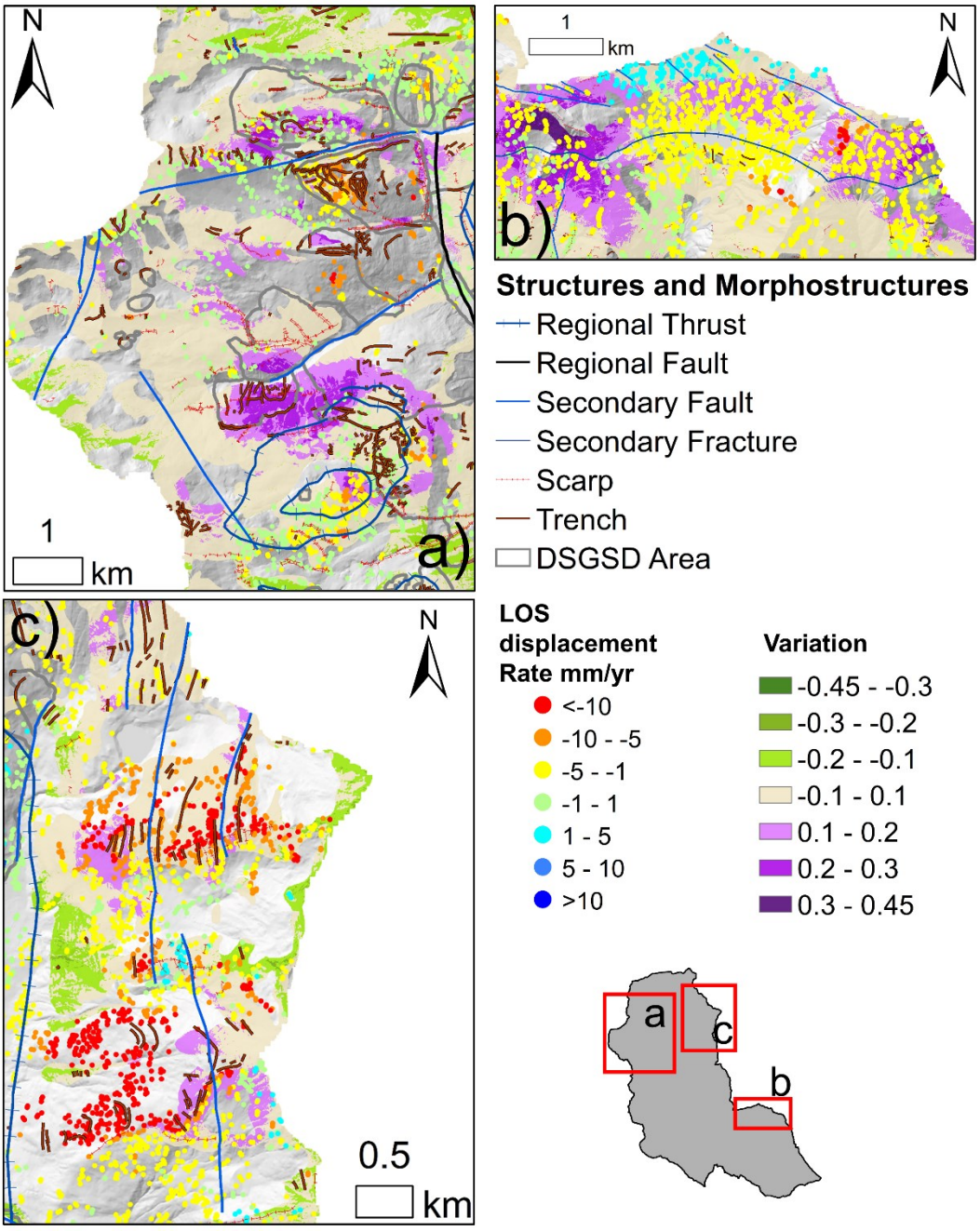


Figure 3.27 GM-TOPO variation map with superimposed morpho-structures and PS/DS derived from inSAR a) Febbraro and Vamlera DSGSDs b) regional thrust area close to the Acquafraggia Lake c) Mount Emet and Mater.

LOS line. Mean annual velocities can overcome 10mm/y along some morpho-structures such as trenches, counterscarps, and structural lineaments. Due to the less favourable exposure of the Bregaglia Valley to the satellites, evidence of movement can be detected only on the northern slope, especially on west dipping facets (Figure 3.27b). These present-day deformation patterns are a good clue for the identification of gravitationally active contexts, in which large and small instabilities coexist, further validating the rockfall susceptibility maps.

3.5.2.4 Uncertainties related to the inventory

The GM and GM+GEO models constructed using the IFFI inventory as input showed a predictive performance in terms of mean AUROC for training and test of 0.77 and 0.75, respectively. Despite these quite good performances, physically distorted, unrealistic relationships between the Jv and infiltration density predictors and susceptibility (Figure 3.28a and Figure 3.28b) were observed. The smoothing functions showed a behaviour opposite to the expected one. Limiting the training set to the IFFI database affected also the relationships between rockfall susceptibility and lithology. The orthogneiss category (Figure 3.28c) became the least susceptible lithology within this model, contrary to the results obtained with the complete inventory. Moreover, implausible rockfall susceptibility maps were obtained, with the highest susceptibility values grouped mainly near the valley bottoms (Figure 3.28d). Finally, the use of the UNIMI inventory as an independent test set led to rather low AUROC values for both the GM and GM+GEO models, 0.69 and 0.61, respectively.

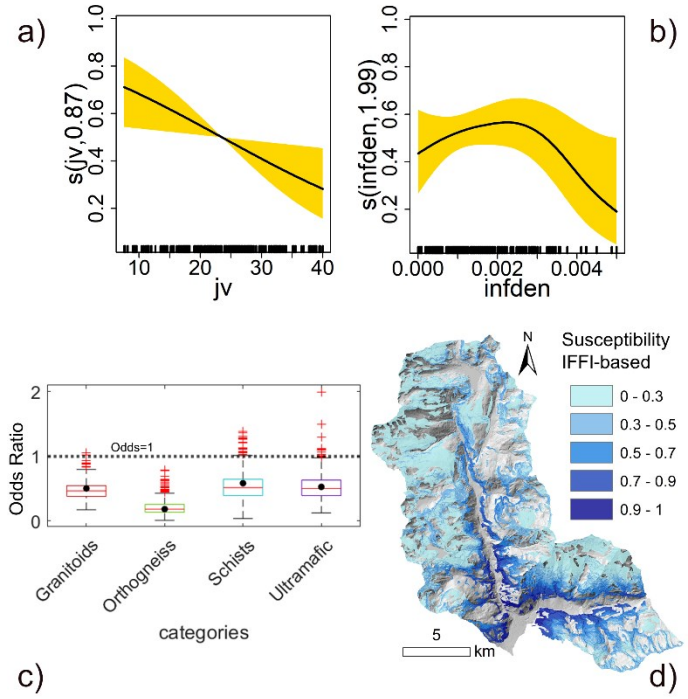


Figure 3.28 smoothing functions for a) Jv and b) infiltration density obtained setting up a geomechanical model based on the IFFI data. The yellow shadow represents confidence bands; c) odds ratios variation among sCV repetitions of the geomechanical-geological model performed on the IFFI inventory. Black dots represent the odds ratios estimated using all IFFI points; d) geomechanical model output map based on the IFFI inventory.

According to the IFFI-based models, the highest fractured rock mass corresponded the lowest rockfall susceptibility level; moreover, the geological peculiar framework of the Gruf Complex seemed not to play a significant role in rockfall occurrence. These distorted outcomes, and implausible susceptibility spatial patterns, revealed a bias in the IFFI inventory. It was prepared with operational purposes and thus linked to damage reported and elements at risk. The additional UNIMI inventory was conversely prepared with geomorphological purposes, independently from the vicinity to vulnerable areas with an accurate and extensive mapping campaign of the area. These results are in line with the findings of Steger et al. (2021) in the Eastern Italian Alps. They observed that, when building models on public reports related inventories, susceptibility outputs clearly reproduced a data collection effects rather than physically driven processes. Similarly, the institutional IFFI inventory was not suitable to produce a reliable rockfall susceptibility map in the study area. Indeed, it cannot satisfactorily and completely capture the dynamics driving the physical processes. This may lead to a misinterpretation of the most critical areas, resulting in challenges for environmental management and planning.

3.5.3. Rockfall susceptibility modelling with Random Forest

3.5.3.1 Hyperparameters and model performance

The best hyperparameters set from the tuning phase was represented by *ntree*=700, *mtry*=9, *node size*=10 and *sample size*=0.63; this set represented the best trade-off between OOB error and AUROC, which showed values of 0.258 and 0.734, respectively. All the 750 combinations of OOB error and AUROC are shown in Figure 3.29.

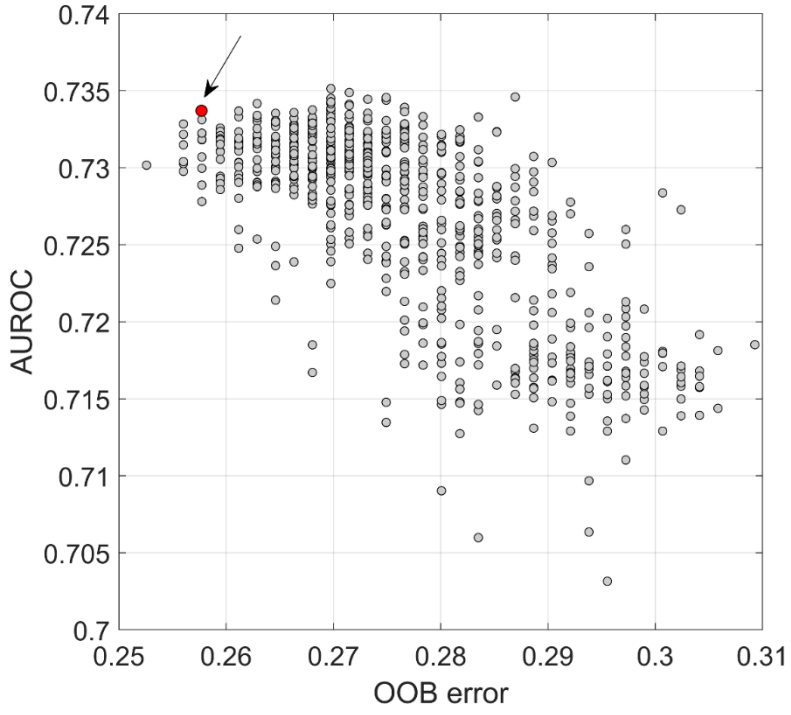


Figure 3.29 OOB error and AUROC for each one of the 750 hyperparameter combinations tested.

It is important to highlight that these two measures are intrinsically different, as the OOB error comes from a sampling with replacement, while the k-fold cross validation AUROC from a sampling without replacement; for this reason, the lowest OOB error may not necessarily correspond to the highest AUROC. Spatial transferability calculated as the IQR of AUROC values on the test set resulted in 0.019. Comparing the performance of the RF model with the GM+GEO model derived from GAM, it is possible to conclude that the two models performed very similarly, with a slightly better AUROC for RF (0.6%). However, the GAM model showed a higher spatial transferability, with a lower IQR of 0.010. This result can be referred to the capability of RF to model more complex interactions than the smoothing functions of GAM; accounting even for relationships between less strong predictors and sub-groups of observations; it follows that some relationships could be more local, thus reducing their transferability on the entire data population.

3.5.3.2 Predictors response relationships

The predictors importance (Figure 3.30a) and physical meaning (Figure 3.30b) derived from the SHAP framework resulted in a good accordance with the GAM smoothing functions and deviance explained, although with some differences. Regarding the topographic predictors, elevation, slope and SWI resulted to be the three most important parameters, followed by profile curvature and northness, whose importance is comparable to each other. SWI acquired more importance than slope in the RF model, while it was at the same level of profile curvature and northness in the GAM model. Plan curvature and eastness were confirmed as the topographic predictors with the lowest importance in the RF model (they were penalized in the GAM model). Summary plots (Figure 3.30b), allowing the analysis of predictors behaviour, also confirmed the same relationships descripted by the GAM smoothing functions.

Regarding the geomechanical predictors, some differences can be observed. While infiltration density and Jv resulted the two most important geomechanical properties, with a direct correlation between feature values and SHAP – coherent with their physical significant behaviour, already observed in the GAM smoothing functions – Keq showed an importance equal to Jv, whereas it was penalized in the GAM model. Moreover, the summary plots (Figure 3.30b) showed a quite clear direct behaviour between medium-high Keq values and SHAP (i.e., high rock mass permeability corresponds to a strong positive impact on susceptibility). Nonetheless, the SHAP behaviour for Keq low-medium values remained quite variable and of difficult interpretation. Among the geomechanical predictors, Wi is placed last in terms of importance. Also, consistently with the GAM model results, the summary plot did not show a clear trend between its values and SHAP so that it could be linked to any physical process. However, it resulted more important than plan curvature, eastness and geology.

Geology showed the same behaviour observed in the GAM model. Orthogneiss (i.e., geo 3 in Figure 3.30) resulted to be the most impacting lithology in terms of susceptibility, consistently with the particular geological-structural framework of the study area, already discussed in Section 3.5.2.3. However, as already seen in the mDD% plots of the GAM model, geology resulted in a small general impact on susceptibility.

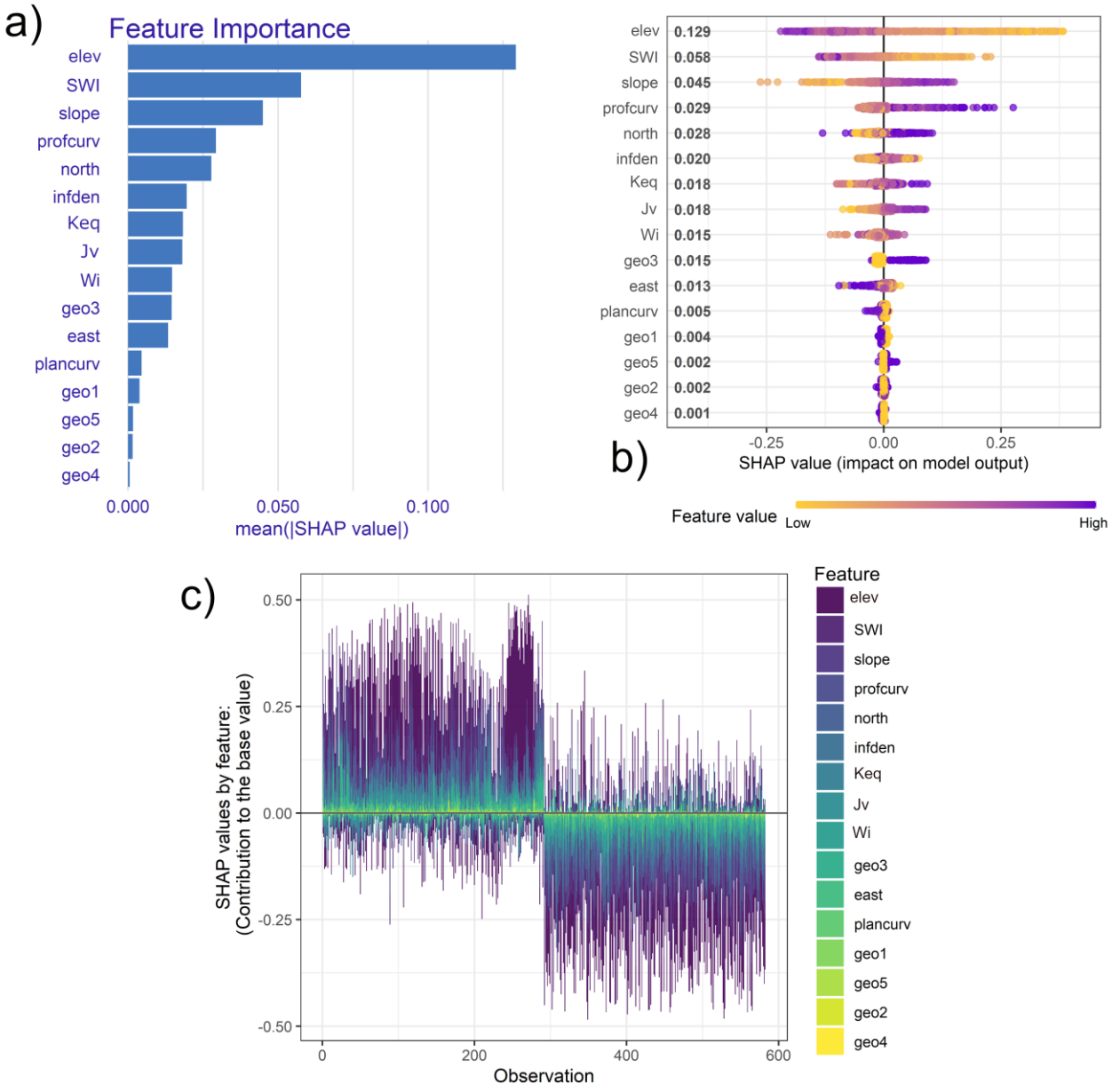


Figure 3.30 a) SHAP feature importance plot; b) SHAP summary plot; c) SHAP force plot. The first half represent rockfall presence, while the second half represent rockfall absence. Abbreviations: SWI=SAGA (i.e., topographic) Wetness Index; profcurv=profile curvature; north=northness; infden=infiltration density; east=eastness; plancurv=planar curvature; geo1=paragneiss; geo2=granitoids; geo3=orthogneiss; geo4=shists geo5=ultramafic rocks.

The force plot in Figure 3.30c confirmed that the most important topographic predictors (elevation, SWI, slope, profile curvature, northness) had the highest impact on the final SHAP value of each observation. Also, it suggested that geomechanical predictors (infiltration density, Keq, Jv) have a medium impact and helped in differentiating, even locally, the Shapely values, while Wi, eastness, planar curvature and geology could be seen generally as noise in the model, with little and localized impact only on the final Shapley value of few observations. An even more intuitive representation of each single predictor behaviour than the summary plot is represented by the dependence plots (Figure 3.31 and Figure 3.32), which allowed a direct comparison with the GAM model CSF plots shown in Figure 3.21. The two series of plots were similar and comparable. For example, elevation showed a secondary peak between 2000 and 2200 m a.s.l. Slope, Jv and SWI showed both a quite linear monotonic behaviour, direct for the former two and inverse for the latter. Northness and profile curvature exhibited a direct increasing behaviour, both showing a not perfectly linear trend for low (i.e., negative) values. Eastness, planar curvature and Wi were essentially characterized by the absence of a clear trend and of a differentiation of predictor values in terms of SHAPs. A brief mention to Keq behaviour is necessary, as it acquired more importance in the RF model. The dependence plot of this predictor seemed to show a cut-off, separating negative and positive impact on susceptibility, at a value approximately of log10(Keq)=-6

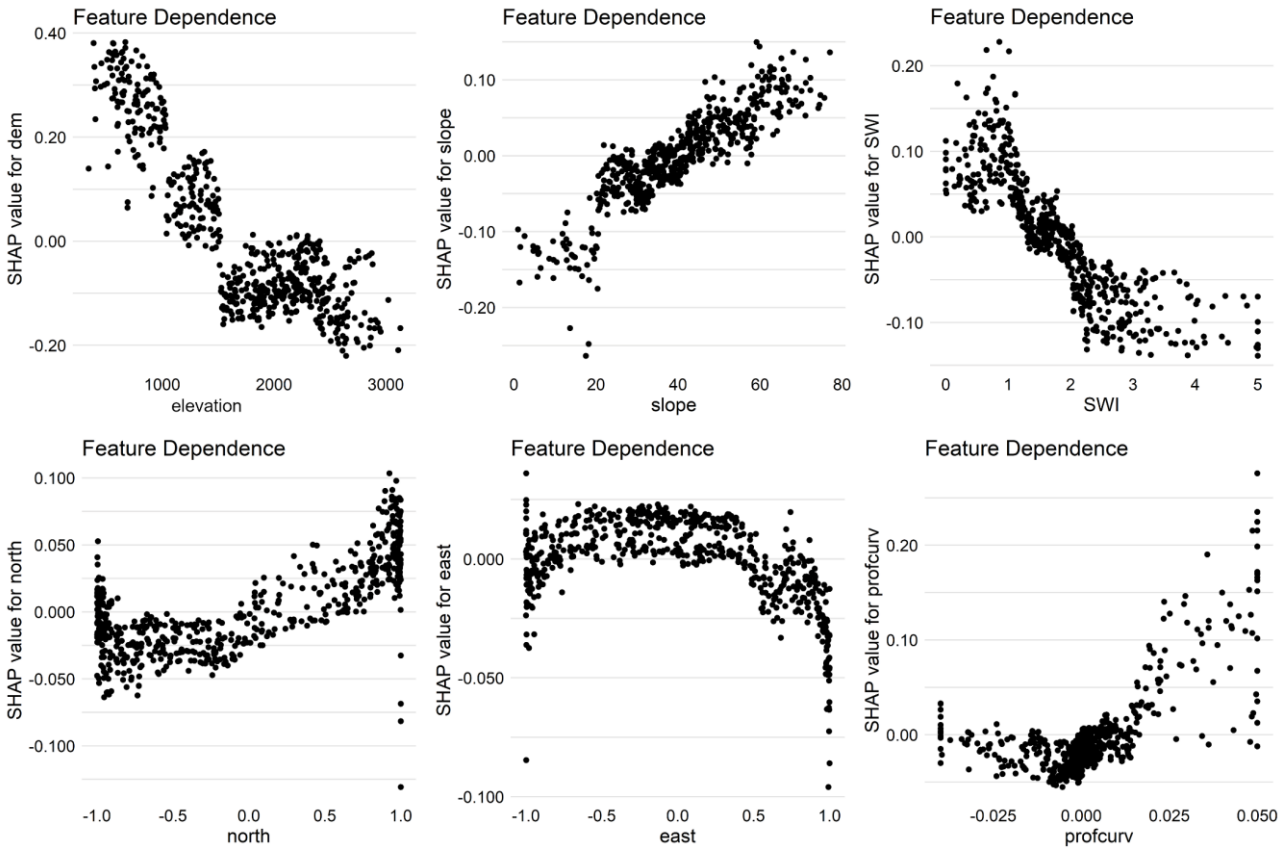


Figure 3.31SHAP dependence plots for continuous predictors (i.e., elevation; slope; SWI; north=northness; east=eastness; profcurv=profile curvature).

(i.e. Keq=10^{-6} m/s). For values lower than the cut-off the dispersion was very high, while for values higher than the cut-off the SHAP impact was positive and with an increasing trend, except for a slightly negative impact peak at a value around log10(Keq)=-3 (i.e. Keq=10^{-3} m/s).

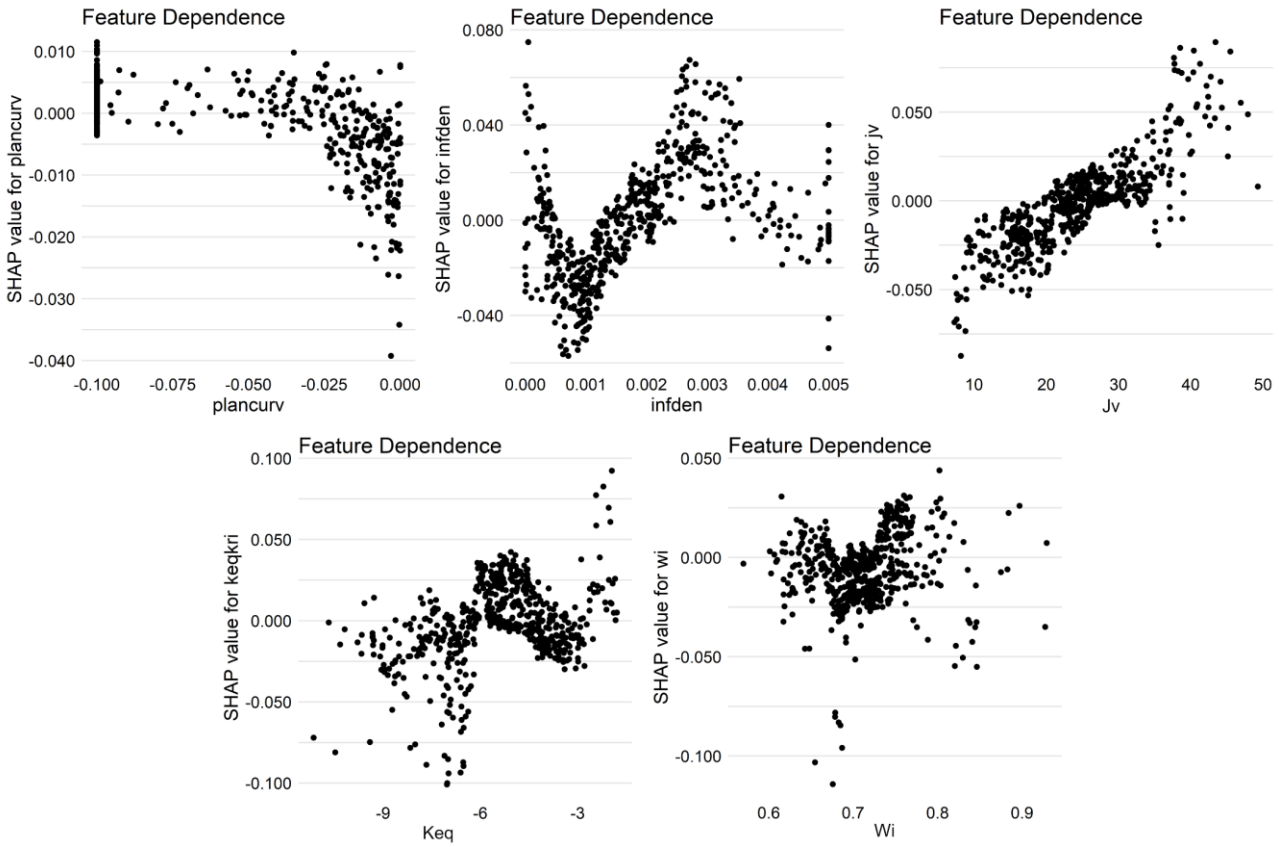


Figure 3.32 SHAP dependence plots for continuous predictor (i.e., plancurv=plan curvature; infden=infiltration density; Jv; Keq; Wi).

To conclude with the SHAP framework analysis, interactions between predictors, with the focus on geomechanical indices and infiltration density, were analysed. It resulted that Jv, Keq and infiltration density had some interactions with elevation, northness and presence/absence of the orthogneiss lithology. Wi did not show any evident interaction. No interaction between the geomechanical predictors themselves were observed. Finally, it is important to highlight, as suggested by Molnar (2019), that SHAP interaction effects are not causal relationships, and these effects might be due to confounding. The principal recognized interactions are shown in Figures 3.33, 3.34, 3.35 for Jv, infiltration density and Keq, respectively. They can be summarized as follows:

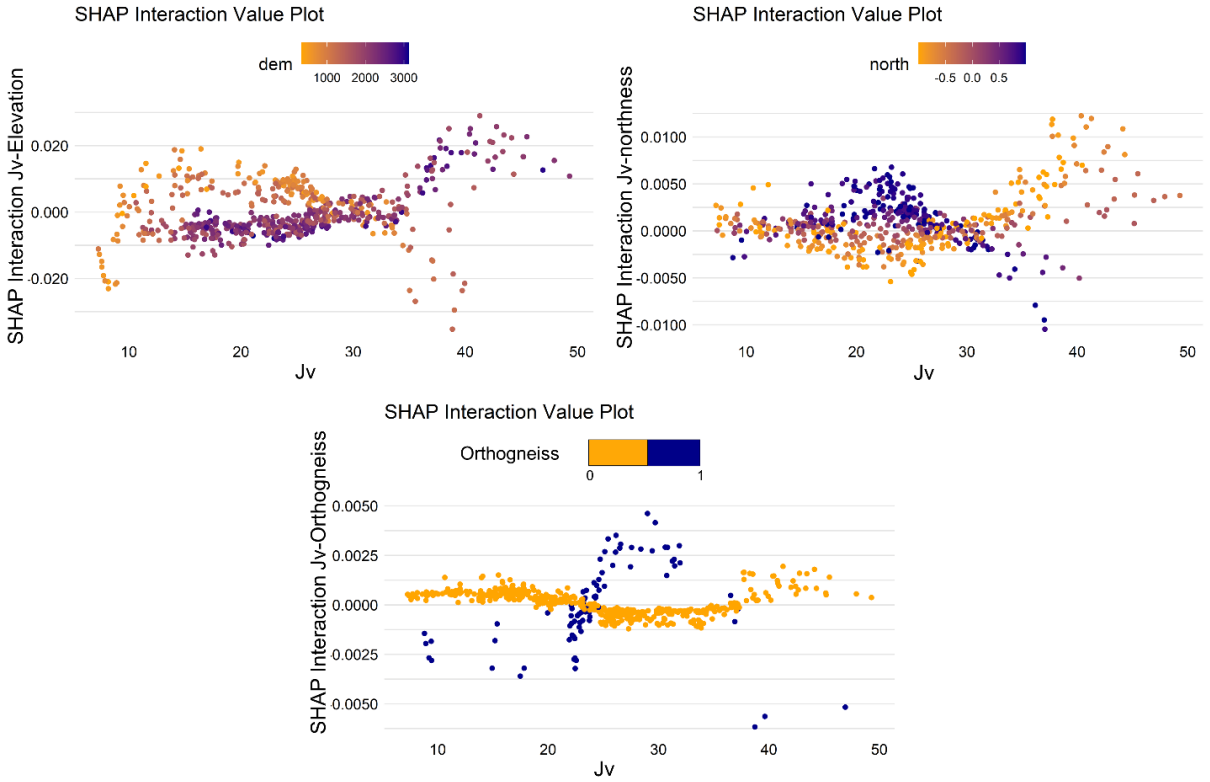


Figure 3.33. SHAP Interactions for Jv with elevation, northness, orthogneiss presence/absence.

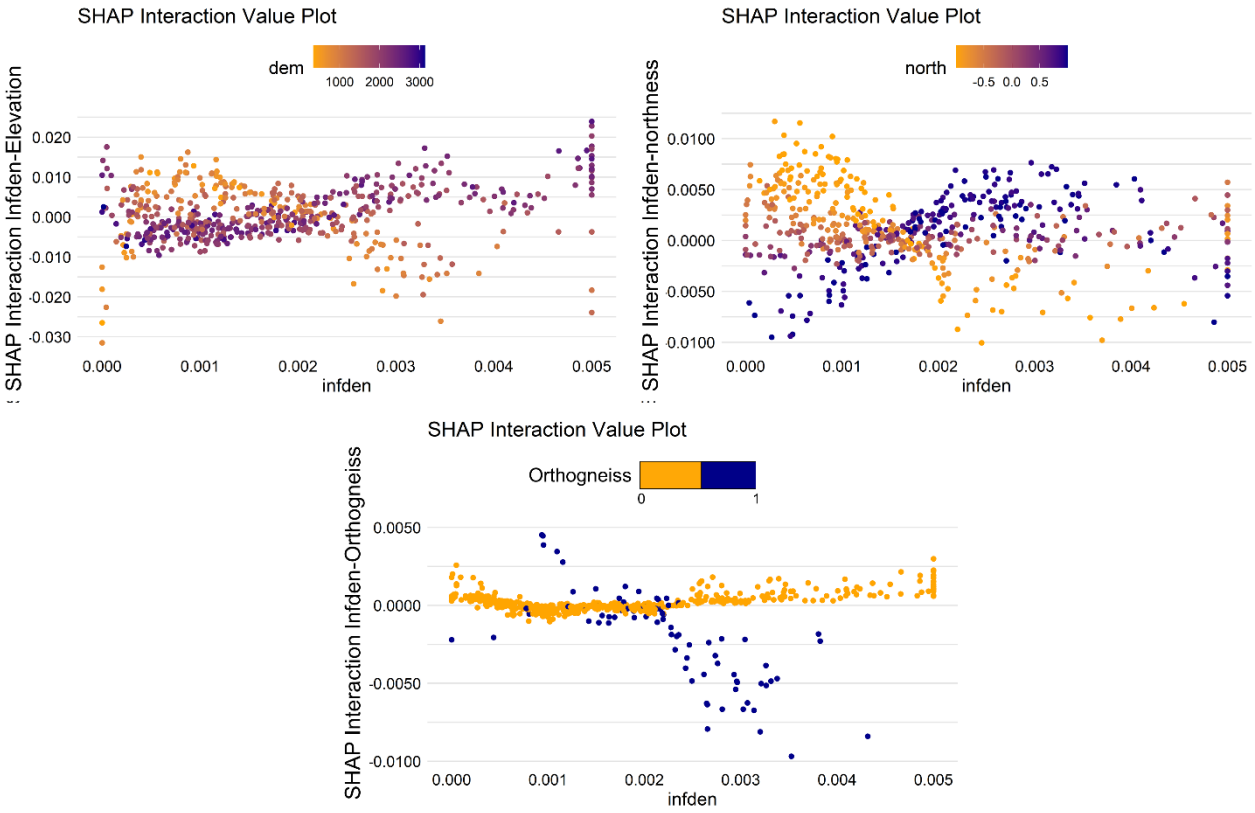


Figure 3.34. SHAP Interactions for Infiltration density with elevation, northness, orthogneiss presence/absence.

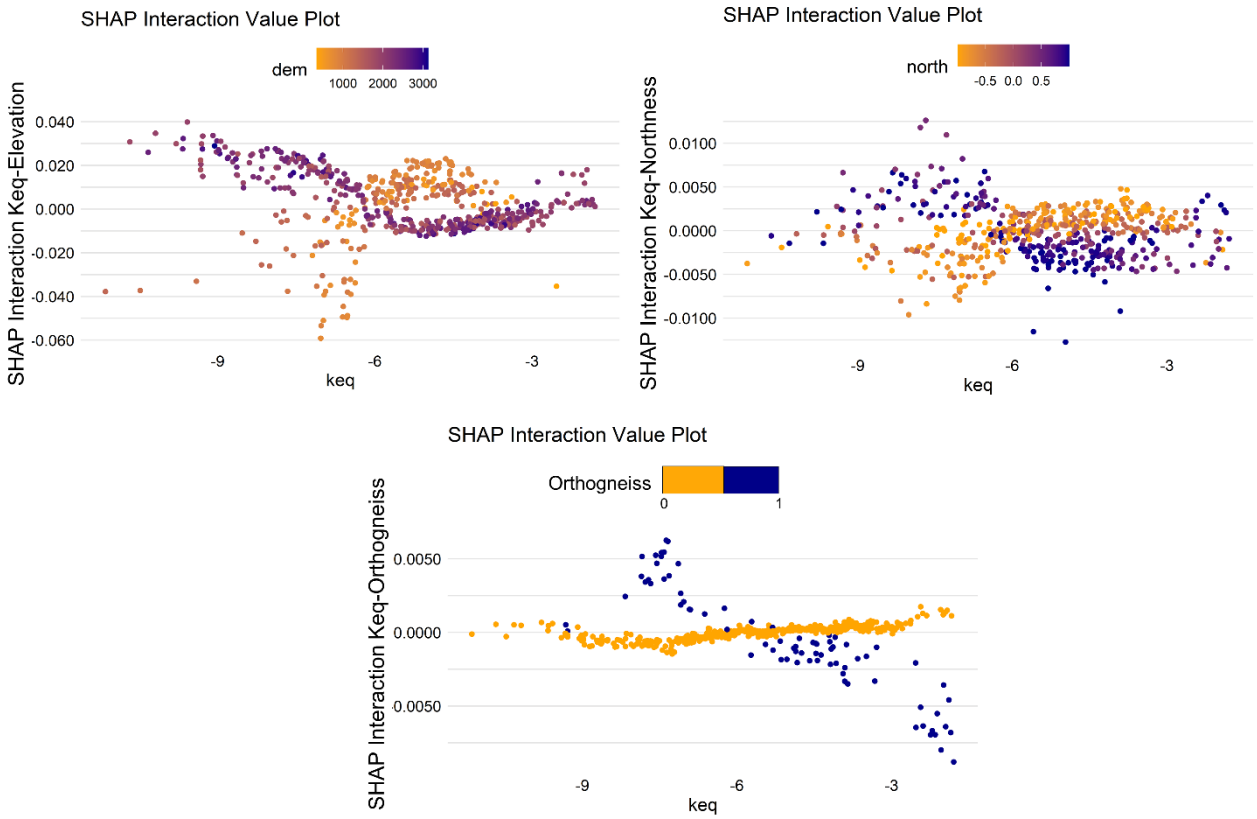


Figure 3.35. SHAP Interactions for Keq with elevation, northness, orthogneiss presence/absence.

(i) Jv (Figure 3.33): a high fracturing grade was more concerning in terms of susceptibility for the RF model at high elevations, in correspondence of south facing slopes and in presence of the orthogneiss lithology (the latter is valid only for Jv values approximately between 25 and 35).

(ii) Infiltration density (Figure 3.34): high infiltration density values were more concerning in terms of susceptibility for the RF model at high elevation, in correspondence of north facing slopes and in absence of the orthogneiss lithology.

(iii) Keq (Figure 3.35): high rock mass permeability values were concerning in terms of susceptibility for the RF model at low elevation, in correspondence of south facing slopes and in absence of the orthogneiss lithology.

3.5.3.3 Uncertainties related to the model selection: comparing GAM and RF

The susceptibility map derived from the RF model is shown in Figure 3.36, and for sake of comparison it was classified with the same classes cut-offs as the GAM map. The spatial pattern was very similar to the GAM model map, but locally subject to less smooth and continuous surfaces, a common graphical aspect related to Machine Learning algorithms (e.g., Goetz et al., 2015). Specifically, this may be due both to the ability of RF to model complex interactions, translatable in more complex predicted surfaces, and to the inclusion of all the predictors in the RF model, which, although with little importance, may be selected in some trees.

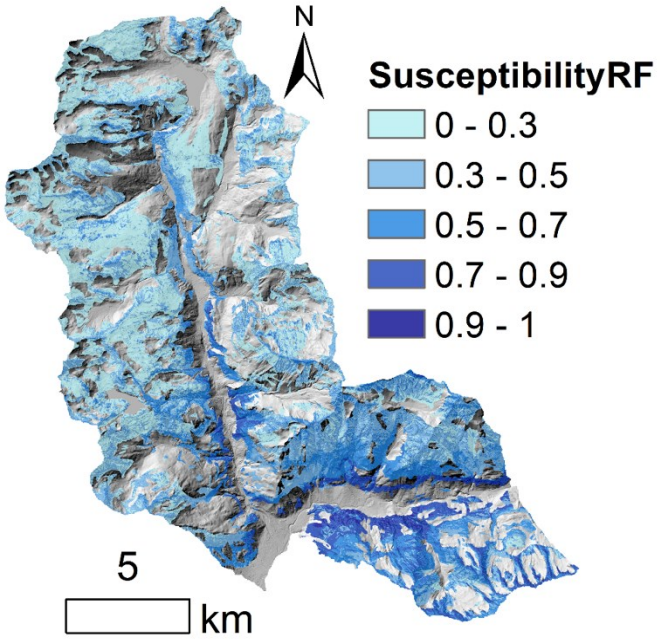


Figure 3.36 Susceptibility map deriving from the application of RF model.

Table 3.18 reports the comparison of area percentages in each susceptibility class between the GAM and the RF model. In comparison to GAM, the RF model distributed almost equally the area between the "low" and "very low" susceptibility classes, while the GAM model predicted more area in the two lowest susceptibility classes. Moreover, the RF model predicted more area in the "medium" susceptibility class and less area in the "high" and "very high" susceptibility classes. The latter has a percentage almost half respect to the one predicted by the GAM model. It can be alleged that the RF model resulted in a slightly more "optimistic" susceptibility map than the GAM model.

Table 3.18 Area percentages in each susceptibility class for the GAM and RF model output maps.

Class	**GAM%**	**RF%**
Very low	43.54	36.60
Low	26.74	34.12
Medium	17.02	20.28
High	11.43	8.37
Very high	1.27	0.64

Figure 3.37 summarizes the comparison of susceptibility spatial pattern changes between the GAM and RF model derived maps. As shown in Figure 3.37a, susceptibility value changes concentrated almost entirely in a range between ±0.3.

The GAM model predicted higher susceptibility than the RF model in Febbraro Valley (NW of the study area) and almost in the entire Bregaglia Valley. Conversely, the RF model predicted a slightly higher susceptibility in the central and NE part of the San Giacomo Valley. However, most of the study area (64%) did not record a susceptibility class changes (Figure 3.37b) with local exceptions characterized

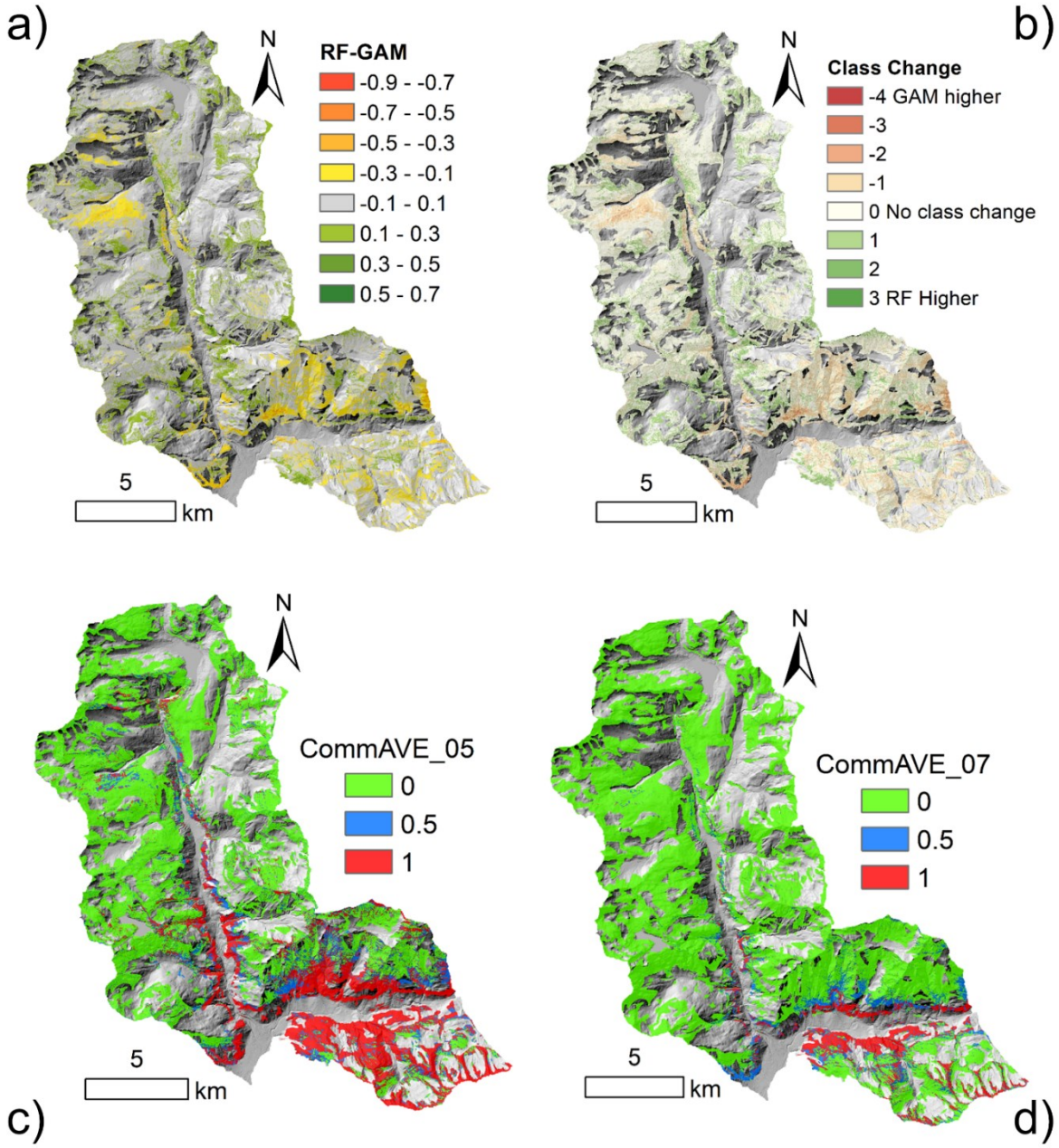


Figure 3.37 a) Map representing the susceptibility difference between the RF and the GAM models output maps. b) Map representing the amount of class change between the RF and the GAM model. c) Averaging Committee ensemble with a 0.5 susceptibility cut-off. d) Averaging Committee ensemble with a 0.7 susceptibility cut-off.

by an upgrade or downgrade of 1 class only (34%). The area affected by 2- or 3-class jumps amounted roughly to the 2% of the total. Concordance and discordance of the two maps were also expressed by the Averaging Committee ensemble procedure with two different cut-offs (Table 3.19).

Table 3.19 Area percentages of GAM and RF models concordance and discordance deriving from the application of the Averaging Committee Ensembles with 0.5 and 0.7 susceptibility cut-offs.

Model Vote	% of Area cut-off 0.5	% of Area cut-off 0.7
0 (Concordance in low susceptibility)	64.02	85.33
0.5 (Discordance)	12.97	7.63
1 (Concordance in high susceptibility)	23.01	7.03

With a cut-off of 0.5 (Figure 3.37c), the discordance between the GAM and the RF model amounted to about 13% and it was located at the middle altitude range in the northern part of Bregaglia Valley and on the right orographic flank of Febbraro Valley. In the latter area, the GAM susceptibility map may be defined as in "favour of safety", in terms of geomorphological significance, assuming a potential high rockfall activity related to the DSGSD framework affecting the slope. On the other hand, RF predicted higher susceptibility than the GAM model in the central part of San Giacomo Valley at low-middle elevations. This is reasonable too, as the valley bottom is bordered by sub-vertical and poorly vegetated cliffs. Both models suggested a high susceptibility to rockfalls for the Bregaglia Valley and for the south-central part of San Giacomo Valley, with the 23% of the study area falling in the highest susceptibility classes. By rising the cut-off to 0.7 (Figure 3.37d), concordance between the two models was higher, with the 7% discordance located again at the middle altitude range in the northern part of Bregaglia Valley.

3.5.4. Combining susceptibility and SAR for operational purposes

The coefficient C map showed that the descending geometry (Figure 3.38a and Figure 3.38b) was more suitable for detecting movements on the slopes located in the eastern flank of San Giacomo Valley (i.e. West facing slopes) and for some west-facing facets located in Bregaglia Valley. Conversely, the ascending geometry (Figure 3.38c) could capture movements in the western flank of San Giacomo Valley (i.e. East facing slopes) and some east-facing facets located in Bregaglia Valley. As expected, San Giacomo Valley, given its N-S direction, was more suitable for SAR data analysis than the E-W Bregaglia Valley.

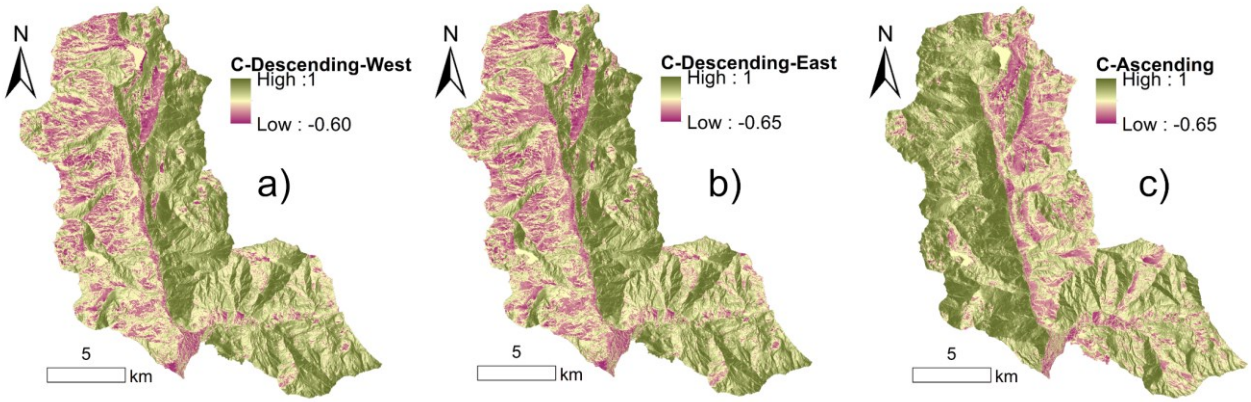


Figure 3.38 Parameter C for a) descending geometry (West tier), b) descending geometr y (East tier), c) ascending geometry.

The Vslope map (Figure 3.39) was no longer related to the single PS/DS but to an average velocity on a 100 m x100 m cells. After post processing by means of the C coefficient, Vslope was measurable on 16.7% of the area in which susceptibility was modelled. This is inherently due to the limitations of the SAR technique, as measurements are affected by vegetation cover, and partly to the E-W direction of Bregaglia Valley, not suitable for a complete detection through Sentinel satellites, especially for the Southern slope. For these reasons, it is important to highlight and to communicate to potential users that the integration procedure was limited to a portion of the territory, and other critical areas may be undetectable by SAR, thus the susceptibility base map remains the most complete tool for risk management purposes.

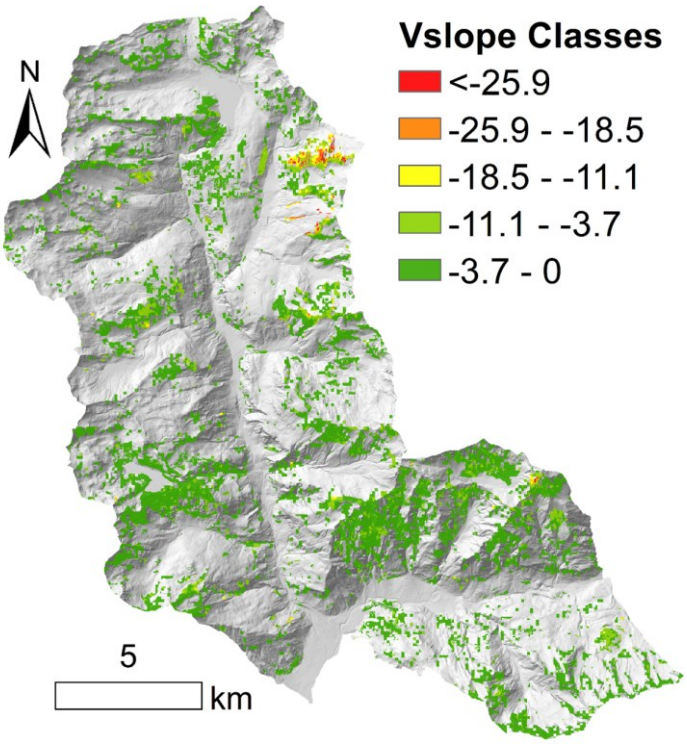


Figure 3.39 Final Vslope layer as the merging of the different acquisition geometries related V-slope

Regarding the SAR Integrated Susceptibility Map, 92.7% of the detectable cells were not subject to class changes due to slope deformation, thus the susceptibility base map remained unvaried. For 5.9%, 0.86%, 0.32% and 0.19% of the detectable cells, changes consisted in an increase of one, two, three and four susceptibility classes due to slope deformation, respectively (Figure 3.40a). In a risk management perspective, these results are quite positive, as these low percentages mean that there are very localized critical areas in which to concentrate time and resources.

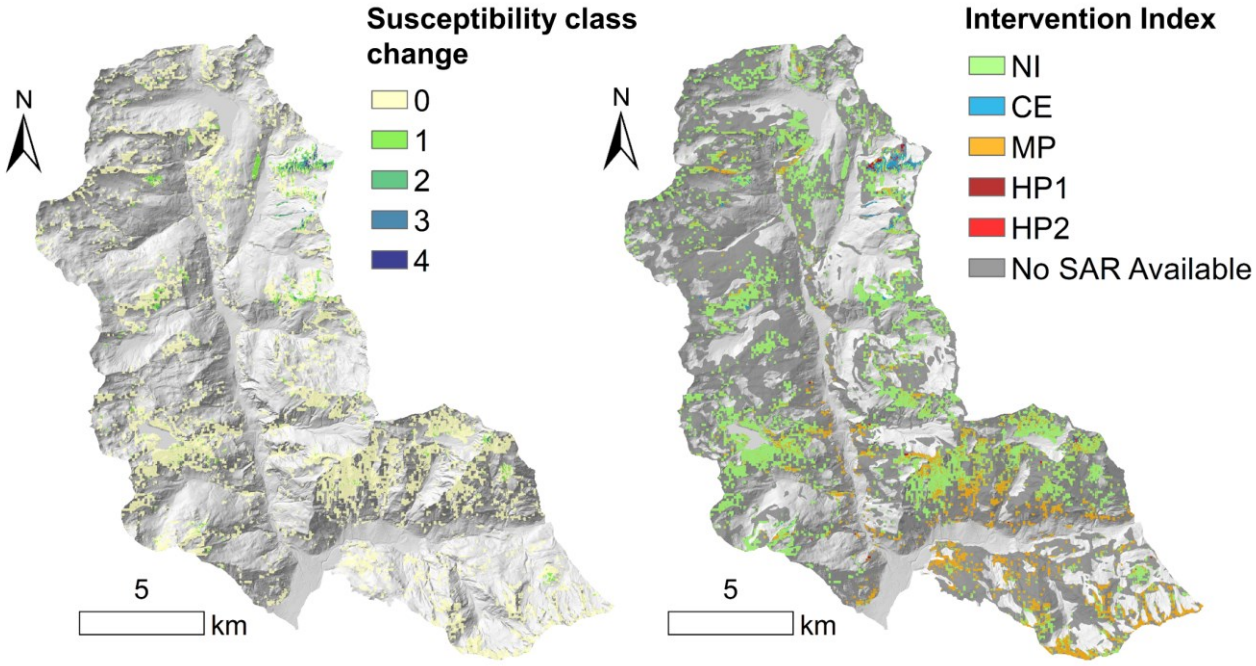


Figure 3.40. a) Susceptibility class change based on the application of the empirical contingency matrix between Vslope and the Mean Ensemble Susceptibility Map; b) Intervention Priority Map.

For what concerned the Intervention Priority map (Figure 3.40b), 74.3% of the detectable cells belonged to the NI class, for which no intervention is required or recommended, as these areas were classified as low susceptibility by the ensemble model and no relevant slope deformation were recorded. Only 1.75% of the detectable cells were classified as CE, which is the category for those areas where the different members of the ensemble model agree in predicting low susceptibility, but where medium to relevant slope deformations were recorded. Causes of this mismatch might be either misclassification errors in the rockfall susceptibility ensemble or the detection of movements related to processes not relevant for rockfall occurrence. For this category, a check of the local processes leading to slope deformation is suggested to understand the source of uncertainty. A percentage of 23% of the detectable cells were classified as MP, i.e., medium priority. In this case, the two susceptibility models (i.e., GAM and RF) agreed in predicting high susceptibility but not relevant slope deformation were recorded in those areas. For this category, careful observations of the slope evolution are recommended. A percentage of 0.35% and 0.36% of the detectable cells were classified as HP-1 (High Priority of type 1) and HP-2 (High

Priority of type 2), respectively. The former category represents those areas in which the ensemble model agrees in predicting high susceptibility and, additionally, medium to consistent slope deformations are recorded. For these areas the uncertainty was therefore quite low and monitoring measures of the involved slopes are highly recommended. The latter category represents areas in which the members of the ensemble model disagreed in predicting susceptibility (i.e., one model predicts high susceptibility and one model low susceptibility), but active slope deformations were recorded. In a land management perspective, for these areas, local analysis, and possibly field work, to enhance the knowledge of the area and to decipher the causes of uncertainty is suggested.

Finally, three areas were selected for a more detailed analysis. They correspond to those where the integration of the susceptibility maps with SAR data showed the majority of changes, and their locations are showed in Figure 3.41.

Area (a) corresponds to the Mount Mater area (already discussed in Section 3.5.2.3): the Mean Ensemble predicted some portions of the slope in “high” susceptibility class, surrounded by other portions in “low” to “medium” susceptibility. This is known as one of the most active slopes of the study area (e.g., Crippa et al., 2020), characterized by the interplay of different processes (e.g., DSGSD, nested secondary rockfalls and rockslides and active rock glaciers). The integration with SAR data led to an overall increment of the susceptibility classes up to four in the SAR Integrated Susceptibility Map, as the entire slope was subject to active deformations. Subsequently, the Intervention Priority Map showed some local areas in category HP-1 or CE. This is an example where the category CE is not present for conceptual errors in the rockfall susceptibility base map (i.e., the Mean ensemble), but rather it may be intended as in “favour of safety” in a risk management perspective. It is indicative of other slope processes or active contexts that are not necessarily directly linked to rockfall occurrence, nevertheless contributing to define a complex deformation pattern worth of attention and in-depth analyses.

Area (b) is located on the eastern flank of the central part of San Giacomo Valley and is rather small. The Averaging Committee Ensemble members agreed in predicting low susceptibility, but some active deformations were detected, leading to a susceptibility increase, although very localized, up to three classes in the SAR Integrated Susceptibility Map. The Intervention Priority Map showed the entire spot in category CE.

Area (c) is located on the northern slope of Bregaglia Valley: in this area, the members of the Averaging Committee Ensemble either disagreed in the prediction or predicted low susceptibility. However, the SAR data used for integration revealed the area as active, leading to an increase of the susceptibility up

to three classes in the SAR Integrated Susceptibility Map. In terms of Intervention Priority Map, the area is either in category HP-2 or CE.

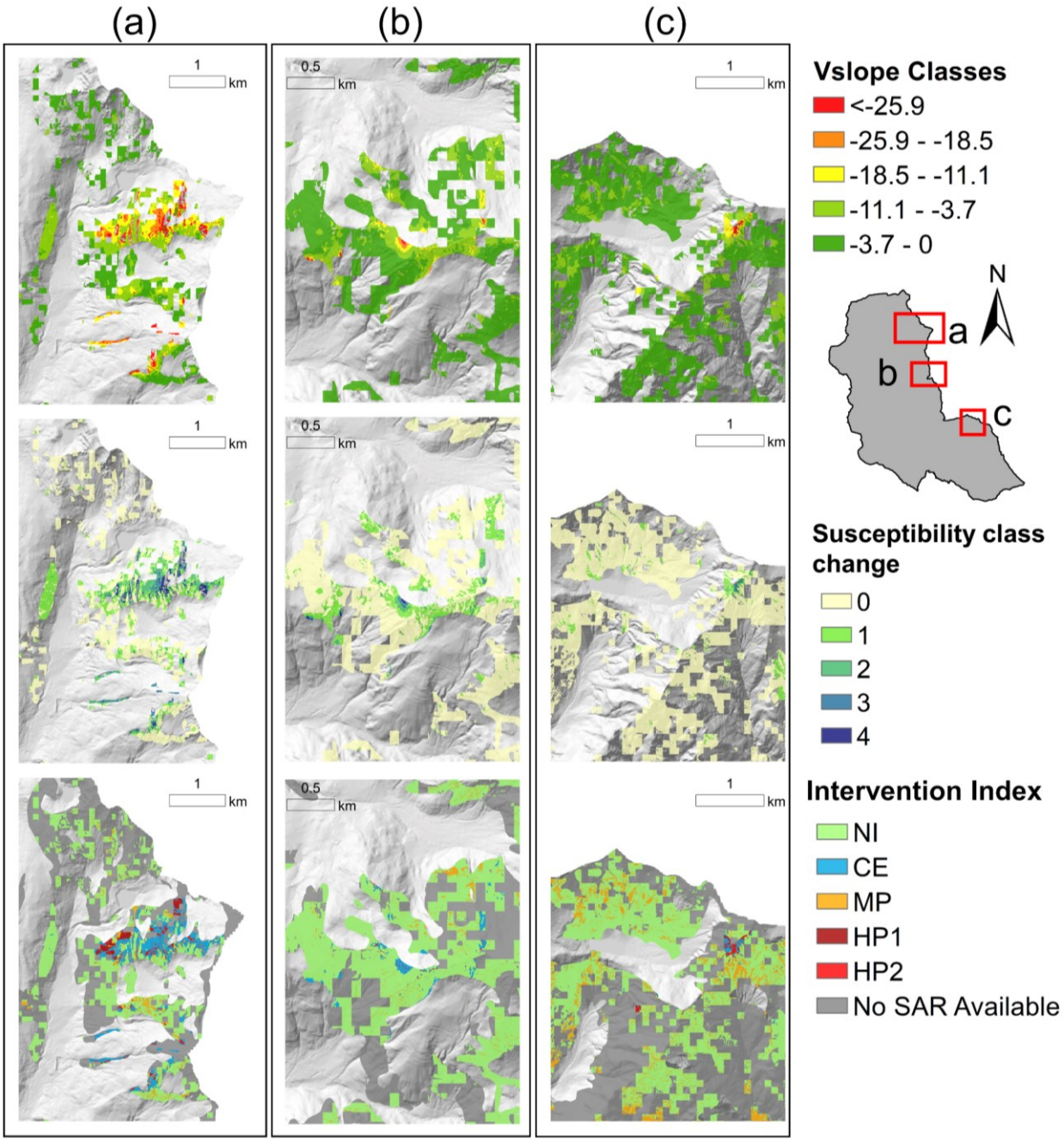


Figure 3.41. Vslope (top), additional susceptibility classes acquired in the SAR Integrated Susceptibility Map (middle) and Intervention Priority map (bottom) of the three selected subareas.

3.6. Discussion and future perspectives

The work was focused on the acquisition and innovative inclusion of geo-environmental data in a rockfall susceptibility assessment process, rather than testing new modelling techniques, as suggested by Reichenbach et al. (2018). Specifically, this case study was devoted to transfer some geomechanical properties (i.e., fracturing degree, rock mass weathering and permeability) relevant at the outcrop scale to a regional scale analysis, traditionally based on topographic predictors, in order to assess the potential added value of this approach in terms of model interpretability.

Two underlying assumptions on geomechanical predictors are worth discussing. The first one is the time-invariance of geomechanical predictors. Even if fracturing, structural framework and weathering grade could be reasonably assumed as static in terms of temporal validity of the susceptibility map, this assumption does not consider the short-term evolution of the investigated systems. This is particularly true when linking weathering to rock slope instabilities. The low influence of weathering grade on rockfall susceptibility obtained in both the GAM and RF models seems to confirm the discussions in Viles (2013), who concluded that the temporal inherent non-linearity of rock mass weathering processes hampers the possibility to make predictions or generalizations about how a change in the weathering component of rock mass would influence its stability. In addition, if on the one hand it was demonstrated by Samia et al. (2017a) that landslides do follow landslides in a path-dependency perspective, on the other hand, when dealing with rock masses, an improvement of the rock mass mechanical conditions may occur after a rockfall. Whether the unstable mass was totally detached, the newly exposed rock surface (which was previously at a higher depth inside the rock wall) could be less weathered (Huisman et al., 2006) and, as joint aperture and joint spacing decrease with depth, less permeable (Snow, 1968; Zhang., 2013). Attempting to introduce these non-linear and feedback processes, however, would require at least a dated multitemporal inventory, combined with a dataset recording both pre-failure and post-failure geomechanical conditions.

Secondly, average representative values for Jv, Wi and Keq were assigned to the point representing the survey. This generalization to the rock mass scale allowed to regionalize Jv and Wi by means of geostatistical technique, already applied in some previous studies aimed at regionalizing some geomechanical properties (Ferrari et al., 2012; 2014; 2019; Pinheiro et al.,2016, 2018; Eivazy et al., 2017; Bajni et al., 2021b). In contrast, the regionalization of Keq was more challenging and led to less encouraging results. This could be linked to the parameters involved in the permeability tensor calculation, which could be very variable even within the same outcrop, which are joint aperture and

Joint Roughness Coefficient (see Coli et al., 2008 for details on the permeability tensor calculation): an average value may not be the best solution. Most probably, these uncertainties in the regionalization of Keq are the reason why it was excluded from the GAM rockfall susceptibility models. Even if this predictor acquired additional importance in the RF model, its behaviour in the SHAP analysis remained quite cryptical, jeopardizing its physical interpretation. More precise and spatially more detailed estimates of Keq than those available may therefore still improve the rockfall susceptibility map in the future. Moreover, in absence of in-situ tests supporting the calibration of Keq (as in Coli et al., 2008 and Oge, 2017), it is more suitable to use this property at outcrop-scale studies, where it is possible and reasonable to capture and represent its own short-scale variability. For basin or regional scale studies, specific research is needed regarding the modification of the permeability tensor equation to include a joint connectivity parameter, which is missing. This gap could be addressed by integrating a synthetic Discrete Fracture Network, starting from the stochastic distribution of rock mass properties or with the emerging application of infrared thermography (Pappalardo, 2018).

The well-established Generalized Additive Models were adopted as the modelling framework. GAMs have already been applied successfully in the landslide susceptibility field (Goetz et al., 2011, 2015; Muenchow et al., 2012; Petschko et al., 2014; Knevels et al., 2020). The algorithm has the advantage of allowing the interpretation of model behaviour, an advisable characteristic for landslide susceptibility models (Brenning, 2012b). With the addition of geomechanical predictors, GAM performance showed a slight enhancement in terms of AUROC. Indeed, it is possible to argue that the time and the cost related to a such intense field acquisition and the following data processing are not worth the effort. However, as pointed out by several studies (Sterlacchini et al., 2011; Goetz et al., 2015; Steger et al., 2016a; Reichenbach et al., 2018; Camera et al., 2021), it is not only a question of model performance in quantitative terms, but it is also a matter of physical and geomorphological plausibility, which is often disregarded. Compared to the topographic model, the geomechanical model related rockfall susceptibility map showed a redistribution of the very high and high susceptibility classes in some relevant and geomorphological plausible contexts. Deepening the knowledge on local geomechanical conditions and including geomechanical predictors in the susceptibility model led to a more process-oriented understanding of the susceptibility spatial patterns, improving the quality of the maps, which is essential for their effective utilization.

Any landslide susceptibility assessment has a level of uncertainty coming from different sources; this is a topic already addressed in different ways by several authors and considered crucial (Guzzetti et al., 2006a; Rossi et al., 2010; Petschko et al., 2014). Alongside model plausibility and interpretability, the

other objective of the work was to evaluate and discuss possible sources of uncertainties arising at each step. The principal source of uncertainty in the procedure was inherently linked to the regionalization of the geomechanical predictors, starting from point data and therefore propagating interpolation errors into the susceptibility model. On the one hand, the dataset was acquired and compiled by different surveyors, therefore the study includes unquantified uncertainties related to subjectivity during field work. On the other hand, each predictor was presented and discussed according to its regionalization performance and error metrics, giving a quantifiable indication of the reliability of the associated predictors maps. Moreover, the quantification of the kriging variance associated to Jv was used to optimize the new geomechanical survey location by means of the SSA algorithm. It was demonstrated that the procedure adopted managed to successfully reduce the errors related to Jv regionalization (and intrinsically the uncertainties) by comparing kriging variance and MESS scores before and after the field survey.

The rockfall inventory used in this study comes from the integration of the institutional IFFI inventory and the one derived from several years of detailed geomorphological field mapping. On the one hand, the integration of the two sources had the aim of reducing the epistemic uncertainty related to inventory completeness. On the other hand, choosing rockfall point sources as both the scarp/crown bisector (for the 83.5% of the inventory) and the highest point of the deposit (for the 16.5% of the inventory) may introduce an error due to lack of positional accuracy on an inventory portion. However, considering the investigated instability - i.e., rockfalls - both the scarp and the rock mass at the top of the deposit could be considered suitable to represent the rock wall geomechanical and morphological characteristics. Indeed, some studies regarding inventory positional accuracy highlighted that using either the scarp and a random point inside the landslide polygon resulted in similar and comparable results in terms of susceptibility model performance (e.g., Petschko et al., 2013, 2016). Consequently, it is reasonable to consider this combined inventory acceptable, although the source of a possible error (and of the inherent uncertainty) is acknowledged (Malamud et al., 2004; Petschko et al., 2014; Steger et al., 2016b). Several authors discussed the importance of a reliable landslide inventory as the foundation underlying susceptibility and hazard assessments (Guzzetti et al., 2012; Steger et al., 2016b). It was demonstrated that the use of the IFFI dataset alone to train the GAM model led to physically implausible relationships between geomechanical predictors as well as geology and susceptibility values, consequently returning implausible susceptibility maps. This is linked to how the IFFI inventory was compiled in the study area i.e., in relation to elements at risk and damages reported by the administration in charge and thus not fully representing the plethora of possible environmental combinations leading to rockfall occurrence. These results agree with the findings of Steger et al. (2016a), who underlined the necessity of supporting quantitative performance measures with geomorphological plausibility and that even strongly biased

inventories could result in high quantitative model performances, especially if linked to confounding predictors (in this study, vicinity to infrastructure, expressed by elevation).

Another common procedure adopted to reduce uncertainties in susceptibility modelling is to apply more than one modelling algorithm (e.g., Pham et al., 2019, 2020, Althuwaynee et al., 2014; Steger et al.,2016b; Goetz et al., 2015; Abedini, 2019; Chen W. et al., 2019, 2020; Pourghasemi and Rahmati, 2018). Two modelling approaches were thus applied to this case study, i.e., Generalized Additive Models and Random Forest. Although RF is a commonly used model in landslide susceptibility, some researchers expressed concerns regarding both its output interpretability and readability (Brenning, 2005, 2012b, Goetz et al., 2015). To overcome these issues, model interpretation was carried out through the application of the SHAP framework (Lundberg and Lee, 2017), frequently employed in other research fields but only once int the landslide susceptibility field by Can et al., 2021 (at least at the time of the present thesis). One of the most important results was that SHAP values confirmed both the behaviour of geo-environmental predictors already revealed by the GAM model and their relative importance in predicting rockfall susceptibility. This outcome added reliability to the physical behaviour of the geo-environmental predictors in the susceptibility analysis. RF quantitative performance resulted to be comparable to GAM, confirming that there is not a "correct" model in a suite of competing models in quantitative terms (Elith et al., 2002; Huabin et al., 2005; Chacón et al., 2006; Reichenbach et al., 2018). Thus, the real challenge consists in maximizing the physical and geo-morphological reliability of the resulting outputs (Sterlacchini et al., 2011; Triglia et al., 2013; Goetz et al., 2015; Steger et al.,2016b).

On the one hand the susceptibility map derived from the GAM model had a reasonable geomorphological plausibility, while the RF model, although producing a very similar map, resulted in a slightly more "optimistic" prediction in some complex and critical (in terms of rockfall activity) geo-structural contexts. On the other hand, due to the capacity of RF to handle complex relationships between predictors, some additional insights regarding reasons leading to high/low rockfall susceptibility were unravelled. In particular, Keq acquired more importance in the RF model than in the GAM model and, despite of difficult physical interpretation, showed a possible cut-off value separating negative and positive impacts on susceptibility. Furthermore, the SHAP interaction plots revealed some links between the geomechanical predictors with elevation, northness and presence/absence of the orthogneiss lithology.

Besides the confirmation of geo-environmental predictors behaviour, the integration of different models in an ensemble brought advantages in terms of uncertainty reduction (Rossi et al., 2010; Di Napoli et

al., 2020; Chen W. et al., 2018; Youssef et al., 2015; Choi et al., 2012; Andan et al., 2020; Rossi and Reichenbach, 2016). The Averaging Committee Ensemble revealed to be a useful tool in a risk management perspective for many reasons (Di Napoli et al., 2020). Firstly, high susceptibility areas individuated by an ensemble of models are subject to a lower uncertainty, helping in prioritizing interventions or monitoring. The same applies for low susceptibility areas. Conversely, in case of ensemble members disagreement (high uncertainty), this approach could help to plan further analyses, field work and additional investigations.

Finally, satellite derived ground deformation data were used both in a qualitative and quantitative way in support of the rockfall susceptibility modelling process. Possible limitations to the use of SAR data in relation to rockfalls is due to the technique itself in terms radar wavelength, revisiting time and the spatial density of measurement points, which usually prevent their utilization for fast moving landslides (Ciampalini et al., 2016). However, slower ground deformations (detectable by SAR), such as DSGSD may reveal critical active processes (e.g., hydrogeological circulation, creep, neotectonics) and are frequently associated to the nucleation of secondary and shallower instabilities (Carrara et al., 1991; Crosta et al., 2013; Crippa et al., 2020). Satellite SAR data were used both to explore the physical plausibility of the GAM derived susceptibility maps and to integrate and verify rockfall susceptibility Ensembles with slope present activity, developing products (i.e., SAR Integrated Susceptibility Map and Intervention Priority map) with land and risk management purposes. The former comparison confirmed that some of the predicted most susceptible areas are actually active. The latter helped in delimiting some critical areas where monitoring activity or additional analysis are required, giving an operative connotation to the model ensembles-related uncertainties.

Chapter 4

Aosta Valley Case Study

4.1. Research question and specific objectives

In mountainous environments the occurrence of slope instability is mostly driven by climate-related processes (e.g., intense rainfall events, prolonged precipitation periods, freeze-thaw cycles and temperature fluctuations above 0°, snow dynamics), which are extremely variable over space and time (Ravanel and Deline, 2010; Ravanel et al., 2013; Collins and Stock, 2016; Paranunzio et al., 2019; Nigrelli et al., 2018; Scavia et al 2020; Camera et al., 2021). In the Alpine region their effects are expected to exacerbate under climate change, modifying the response of mountainous environments to instability (Crozier, 2010, Stoffel et al., 2014; Gobiet et al., 2014; Beniston et al., 2018).

Rainfall processes related to rock slope instability include: (i) rising of water pressure in joints with a consequent reduction of shear strength; (ii) joint clayey infilling swelling, dissolution, and leaching with a consequent reduction in joint cohesion; (iii) reduction of the strength of rock bridges within rock masses. These processes could act at different time scales, depending on the intensity and duration of the rainfall events. Together with precipitation, the action of snow accumulation and melting influences rockfall occurrence, being both an immediate and cumulative source of water impacting rock masses, similarly to rainfall (Lucas et al., 2020; Stumvoll et al., 2020; Subramanian et al., 2020). In particular, the intermittence between snowfall on a rock surface close to 0 °C and daytime insolation may lead to rising moisture content, implying the rapid infiltration of snowmelt water in the bedrock discontinuities and pores (Matsuoka, 2019). Moreover, the permanence of snow may isolate bedrock from external temperatures, influencing the occurrence of freeze-thaw cycles and the evolution and onset of permafrost (Magnin et al., 2017).

Temperature variations can affect rock mass properties as well. Given partially saturated joints, temperature can deteriorate rock masses through transitions across 0 °C (i.e. freeze-thaw cycles). Freeze-thaw cycles affect rock masses through ice formation-melting processes, i.e. frost cracking and frost weathering (Hales et al., 2007; Matsuoka, 2008). Acting on the long-term, freeze-thaw cycles enable ice segregation-induced subcritical cracking, progressively decreasing rock walls strength (Draebing and Krautblatter, 2019) and inducing rock thermal weathering (i.e., repeated expansion and contraction

leading to rock mass fatigue). Even a negative warming can induce crack propagation due to the thermal expansion of ice (i.e., thermal wedging, D'Amato et al., 2016). Furthermore, positive temperatures and gradients both lead to ice melting, and thus to a loss of cohesion. In the light of global warming, an interest in analysing the role of positive temperatures on rockfall occurrence is spreading in the scientific community. In arid and semiarid environments very high positive temperatures (up to 40° degrees) could lead to rock mass exfoliation and fracturing through thermal stress cycles (Bakun-Mazor et al., 2013; Vargas et al., 2013; Collins and Stock, 2016; Collins et al., 2018). Conversely, authors working on alpine rockfalls recognized the role of extremely warm temperatures principally on permafrost degradation and snow melt dynamics, which lead to an increasing water availability and circulation in rock joints (Allen and Huggel, 2013; Bodin et al, 2015; Paranunzio et al., 2016; Nigrelli et al., 2018). Nonetheless, in alpine environments, thermal stress weathering is principally related to the occurrence of the transition from above to below 0 °C and *vice versa* (Nigrelli et al., 2018).

The wet-dry cycles are another important rockfall climate-related preparatory factor that is often neglected at the slope scale, but frequently investigated at the lab scale (Van der Hoven et al. 2003; Torres-Suarez et al. 2014; Zhou et al. 2017; Yang et al., 2018, 2019). Water can interact with rocks through pores, fissures and microdefects, resulting in changes in several micro-structural characteristics. The rock microstructure damages have a macroscopic manifestation in the deterioration of mechanical properties and the increase in deformation of rock masses. The damage caused to rock masses by multiple cycles is gradually accumulated and irreversible. The weakening effect of these cycles on rocks is often stronger than that of being soaked in water for a long time and thus it seriously influences the long-term stability of rock masses (Yang, 2018).

Previous studies focusing on the interaction between rockfalls and meteorological events (e.g., Matsuoka, 2019; Paranunzio et al., 2015, 2016, 2019) highlighted the extreme variability of these relationships. Authors mostly investigated short term effects related to rainfall (i.e., rainfall intensity and amount at the sub-daily, daily or weekly scale) or long-term effects related to freeze-thaw cycles and temperature extremes (i.e., at the monthly, quarterly or longer period scale). Their findings testify that both the short-term and long-term effects should not be neglected, representing triggering and preparatory factors for rockfall instability, respectively. Triggering factors are immediate causes acting directly, while preparatory factors are linked to a slow cumulative effect, requiring a higher amount of time to induce a major consequence (Gunzburger et al., 2005).

In the Japanese Alps, Matsuoka (2019) identified five key climate-related processes controlling rockfall occurrence: (i) summer and early autumn heavy rainfall events; (ii) lighter, repeated rainfall events

leading to a moisture increase followed by freeze-thaw cycles, occurring primarily in spring and autumn; (iii) same as (ii) but with snowfall replacing rainfall, in winter and early spring; (iv) seasonal thawing after deep winter freezing; (v) thermal stress induced by large thermal fluctuations (mainly in winter). In the Italian Alps, Paranunzio et al. (2015, 2016, 2019) proposed an approach to identify climate anomalies associated with the occurrence of rockfalls at high elevations. The approach includes the analysis of temperature, temperature variations and rainfall data by comparing the values observed for these climatic variables before rockfall occurrences with the empirical distribution function of a reference sample. To derive the latter, all the available values recorded for the climate variables in the same period of the year with different temporal aggregations are used. They found different positive and negative anomalies (i.e., values corresponding to the tails of the distribution) associated to rockfall occurrence. In particular, at elevations between 1500 and 2400 m a.s.l., rockfalls occurred mainly in spring and were mostly associated with negative temperature anomalies. At high elevations (above 2400 m a.s.l.), rockfall events concentrated in summer and positive temperature anomalies prevailed as triggering conditions. Only 15% of the rockfalls recorded in the database were associated with exceptional precipitation in the 7–90 days preceding failure. During four years of rockfall measurements by means of nets and traps in the German Alps, Klaukblatter and Moser (2009) found that rockfall intensity is coupled to rainfall intensity only if it exceeds a certain threshold (i.e., 9–13 mm/30 min). Also, they found that the rockfall response to rainfall intensity above the threshold is highly nonlinear. In the French Alps, D'Amato et al. (2016) observed that the highest daily-weekly rockfall frequency of a rock wall near Grenoble occurred during freeze-thaw cycles, especially during thawing periods, but the highest hourly rockfall frequency occurred during intense rainfalls (more than 5 mm h^{-1}). In another study area of France, Delonca et al. (2014) observed that the most influential parameter leading to an increase in rockfall frequency was the cumulated rainfall in the antecedent 3-day period. In Switzerland, Strunden et al. (2015) compared different size groups of rockfall events with environmental factors (e.g., freeze-thaw cycles, temperature, precipitation, and seismicity) using a linear regression with variable lag times between 0 to 6 months. The highest correlation factor was observed for freeze-thaw cycles and rockfall events involving volumes smaller than 1 m^3 with a 2-month delay between temperature extremes and rockfalls. A rather high correlation was discovered between cumulated precipitation and rockfall events for a 4- to 6-month lag time, too. Along a railway section through the Canadian Cordillera, Macciotta et al. (2015) found that the seasonal variation in rockfall frequency was mostly associated with cycles of freezing and thawing during the winter months. They also compared the intensity of precipitation and freeze–thaw cycles for different time-windows against recorded rock falls and used their findings to propose a rockfall hazard chart. In Aosta Valley, Ponziani et al. (2020) analysed the

relationship between debris-flows and the dynamics of freezing level, permafrost temperature and rainfall, in order to account for these hydro-meteorological processes in the regional Early Warning System.

Studies regarding the influence of climate factors allow to design triggering thresholds. A threshold is a curve that may define the conditions that, when reached or exceeded, are likely to trigger landslides. Thresholds can be defined on physical or empirical bases. Intensity-Duration thresholds for landslide initiation are the most common type in literature (Guzzetti et al., 2007 and references therein). They usually have the form of a negative power law:

$$I = c + \alpha D^{\beta} \qquad \text{Eq. 4.1}$$

where I is the rainfall intensity, D is rainfall duration, and c (≥ 0), α, and β are empirical parameters. They are very common for shallow landslides and only in few cases they consider other types of mass movements, including rockfalls. Reasons are the easily detectable relationships between shallow landslides and extreme meteorological events and their occurrence on gentler and less remote slopes in comparison to those affected by rockfall. Furthermore, quantitative thresholds regarding freeze-thaw or wet-dry cycles are very rare in literature. One of the few examples is that proposed by D'Amato et al. (2016), who quantified that rockfall frequency can be multiplied by a factor of 7 during freeze–thaw cycles.

Despite the presented studies and approaches, many authors agree that additional research is needed to improve the understanding and quantification of the relationships between climate processes variability and geo-hydrological hazards. Deciphering the role of climatic factors is a fundamental research challenge that needs to be addressed to elaborate adaptation strategies for the prevention and the mitigation of climate change impacts in mountainous environments (Crozier, 2010, Stoffel et al., 2014, Gariano and Guzzetti, 2016; Nigrelli et al., 2018; Paranunzio et al., 2019). According to some authors (Gariano and Guzzetti, 2016; Reichenbach et al., 2018), it should be the main thread guiding studies on landslide susceptibility, hazard and risk. Nevertheless, focusing on susceptibility modelling, studies including climate related variables are quite rare, with only 2.8% of them including a precipitation-related predictor and only 0.3% including other climatic predictors (Reichenbach et al., 2018). This may be linked to how susceptibility is perceived, and related studies are commonly addressed; indeed, the traditionally used geo-environmental predictors (amongst all, those derived from the DEM) are considered static and, consequently, susceptibility is inherently deemed as stationary (Lombardo et al., 2020). Moreover, it is often assumed that climatic processes are related to the temporal occurrence of

landslides and not to their spatial distribution (Pereira et al., 2012), which is the aim of susceptibility analyses. On the other hand, climate-related processes affecting slope instability are dynamic in time and space, thus synthetic process-related non-stationary variables, adaptable to reveal climate change impacts are necessary (Camera et al., 2021). Furthermore, the validity of the assumption that climate processes would not impact on landslide spatial occurrence is suitable only over small areas where the conditions are homogeneous and not for large areas where different microclimatic conditions may occur (Catani et al., 2013).

In the most recent years, researchers focused on addressing this gap and different approaches were tested, mainly regarding the introduction of rainfall related variables in shallow landslide susceptibility models. The most basic approach is to include precipitation in the form of mean annual rainfall, mean monthly rainfall and rainy days frequency (e.g., Broeckx et al., 2018; Chen and Li, 2020; Fang et al., 2020; Nahayo et al., 2019; Nhu et al., 2020). The flaw in this approach is that it may not capture the weather conditions leading to landslide occurrence, which are widely variable in terms of amount, duration and intensity (Perruccacci et al., 2017). In other works, researchers developed event-based susceptibility studies, introducing a particular intense and prolonged rainfall event summarized as a multiple day-maximum cumulated precipitation variable and exploiting it to model post-event rainfall induced landslides (e.g., Kim et al., 2015; Knevels et al., 2020). A possible shortcoming is that these models are valid only for a limited time frame and a single precipitation event characteristic. Finally, traditionally stationary variable-based susceptibility models are updated for early warning purposes by coupling them with thresholds exceedance (Segoni et al., 2015, 2018) or by multiplying the stationary spatial model with an additional temporal statistical model including antecedent cumulated rainfall and soil saturation degree (Bordoni et al., 2020). Only few researchers attempted to further synthetize rainfall-related variables and formally introduce them in a single modelling phase. As an example, Catani et al., (2013) developed seven variables expressing the return period of a rainstorm characterized by a given total rainfall amount in a given time lapse. More recently, Camera et al. (2021) introduced the annual number of rainfall events with intensity–duration characteristics above a defined threshold and the average number of snow melting events occurring in a hydrological year. The latter is one of the few recent works in the related literature attempting to introduce a climatic variable different from rainfall. The extreme scarcity of works considering climatic predictors different from rainfall and coupling climate with rock mass instabilities in a susceptibility model rather than shallow landslides, represented the main research idea at the basis of this part of this doctoral study.

Another open question in landslide susceptibility analysis relates to the physical and geomorphic plausibility of the optimized statistical model (Steger et al., 2016b, 2021; Camera et al., 2021), which is nonetheless strictly linked to the characteristics of the landslide inventory available and of possible related bias and inaccuracy (Steger et al.,2016a, 2017, 2021; Bajni et al., 2021b). The consistency and coherency of process-related predictors behaviour in the model with their physical meaning and role on landslide occurrence is absolutely crucial when dealing with climatic processes (Camera et al., 2021). Indeed, focusing only on quantitative model performance and predictors importance rather than on predictors physical behaviour and inventory related issues may lead to severe errors in the risk management process (Carrara et al., 1991; Reichenbach et al., 2018; Steger et al., 2021; Bajni et al; 2021b). Due to the evidence that models with very similar performances may produce very different susceptibility maps (Sterlacchini et al., 2011; Triglia et al., 2013; Goetz et al., 2015), a process-driven understanding and awareness is crucial in the selection of the best, most suitable model in a multitude of possible, equal-performing ones.

To summarize, the overarching goal for the Aosta Valley Case Study was to test the influence of climate-related spatially distributed predictors on rockfall susceptibility in an Alpine environment. The study focused over the Mountain Communities of Mont Cervin and Mont Emilius (central part of Aosta Valley, Western Italian Alps), where a large historical rockfall inventory and an extensive, multi-variable meteorological dataset are available for the period 1990-2018 (2020) (i.e., a three-decade period allowing the calculation of Climate Normals as defined by the World Meteorological Organization, WMO 1989, 2007). Specific objectives were: (i) the recognition of climatic processes influencing rockfall occurrence and consequent development of empirical critical thresholds; (ii) the conversion of critical thresholds in potentially dynamic summary climatic predictors for rockfall susceptibility modelling; (iii) the set-up of rockfall susceptibility models by means of Generalized Additive Models (GAM) in which the physical plausibility of the outputs is guaranteed.

4.2. Study area

Aosta Valley is located in the North-western Italian Alps. It is the smallest Italian region, with an area of 3262 km^2 (Figure 4.1a). The region represents one of the main alpine valley systems and comprises some of the highest mountain peaks of the Alps, among which Mont Blanc (4810 m a.s.l.), Monte Rosa (4634 m a.s.l.), Matterhorn (4478 m a.s.l.) and Grand Paradiso (4061 m a.s.l.).

The Aosta Valley geo-structural setting is the outcome of the collision of three main continental domains, namely the European palaeomargin, the Briançonnais microcontinent and the Adriatic palaeomargin, originally separated by the Valaisan and the Piemonte-Ligurian oceans; it was extensively analysed in literature as it is an example of a complete section of the alpine orogenic prism (Argand, 1911; Dal Piaz,

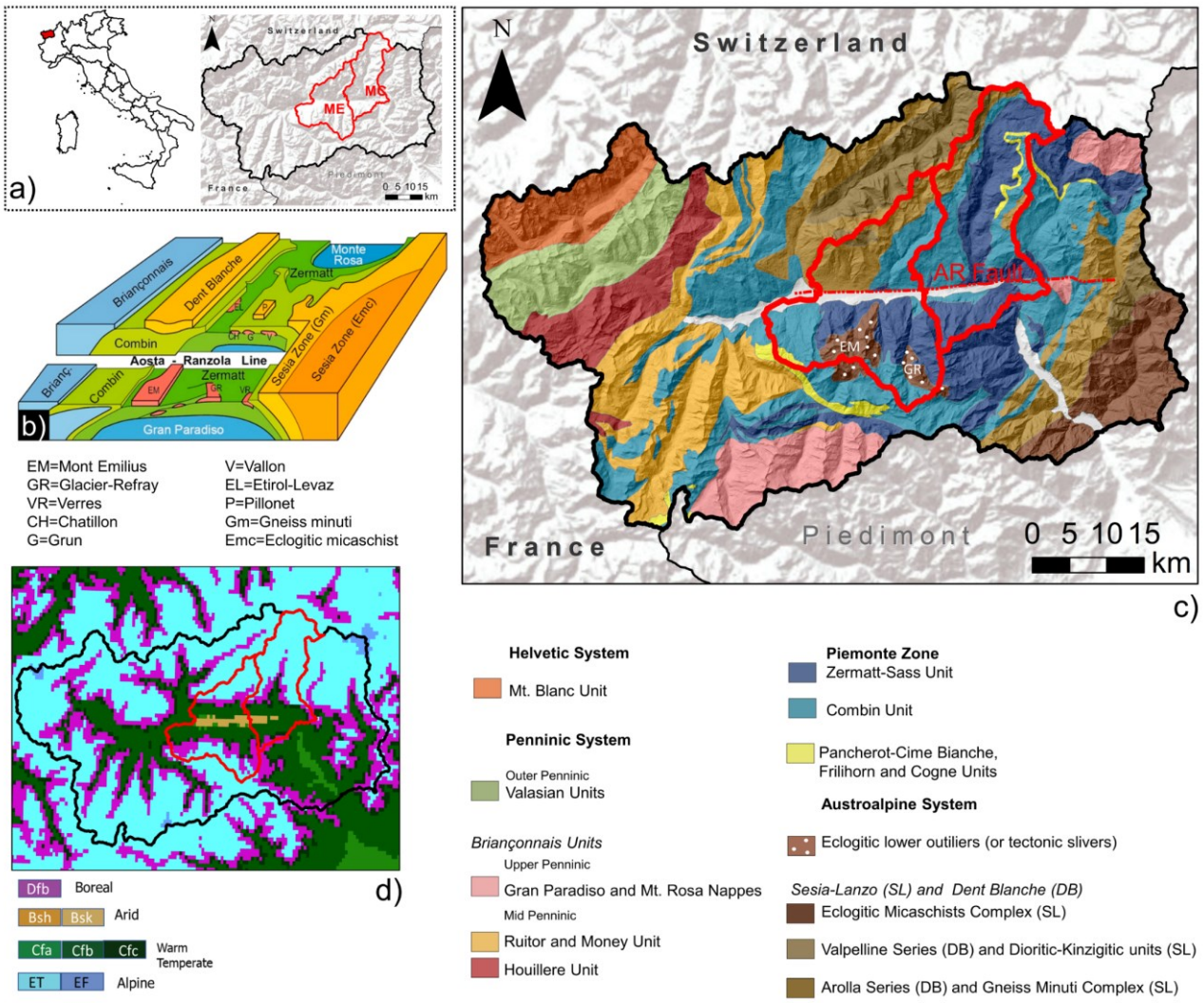


Figure 4.1 a) Aosta Valley region and study area location (red borders). b) Structural sketch of the study area (from Tartarotti et al., 2019. c) Geos-structural map of Aosta Valley (modified from Bigi et al., 1990 and Ellero & Loprieno, 2017). d) Aosta valley Koppen's Climatic classification (Map available at http://koeppen-geiger.vu-wien.ac.at/alps.htm, Rubel et al., 2017).

2001; Giusti et al, 2004; Ballevre et al, 2015; Ellero and Loprieno, 2017 and references. therein). The territory extends through the Europe-verging structural domains of the Western Alps, crossing the complex pile of continental nappes and minor ophiolitic leaves, characterized by ductile deformation and a differential subduction related metamorphism, ranging from the blueshist to the eclogitic facies, and locally ultra-high conditions (Frey et al., 1999), indicating that each nappe had been subducted at different depths and exhumated independently (Bistacchi et al., 2001). From top to bottom (Figure 4.1b, c), it consists of (Bistacchi et al., 2001):

- The Adria-derived Austroalpine system, consisting of pre-alpine granitoids, paragneisses, granulites and michashists in eclogitic facies of the Sesia-Lanzo inlier and a number of upper (Dent Blanche s.s., M. Mary, Pillonet, Mont Cervin, Rosain) and lower (M. Emilius, Glacier-Rafray, Chatillon-Saint Vincent, Etirol-Levaz) outliers, with a blueshist and eclogitic metamorphic imprint respectively.
- The ophiolitic Piemonte zone, including the structural composite upper Combin zone and lower Zermatt-Saas zone. The Zermatt-Saas Zone comprises dominant serpentinized peridotites associated with a metasedimentary cover consisting of metaquartzites, marble, and calcshists, all showing effects of high-pressure subduction-related Alpine metamorphism (i.e., eclogitic facies prevalently). The Combin Zone records the blueshist facies and consists of carbonate and terrigeneous flysch-type calcshists hosting tectonic sheets of mafic and ultramafic rocks (Tartarotti et al, 2019).
- The Europe-derived continental nappes including: (i) the internal Penninic units of Monte Rosa, Arcesa-Brusson and Grand Paradiso, with an eclogitic metamorphic imprint and the mid-Penninic multinappe system (Houliere and Ruitor Units) with a low-pressure metamorphic overprint under blueshist or greenshist facies conditions, both deriving from the Briançonnais domain (Ellero and Loprieno, 2017); (iii) the outer Penninic flysch-dominated Valais zone, recording high pressure metamorphism. The SE-dipping Frontal Pennine Fault separates all these units, representing the axial sector of the alpine belt, from the Helvetic Mont Blanc basement, composed of pre-alpine amphibolite facies metagranitoids and migmatites

Following the Alpine orogenesis, this complex overlapping structure is nevertheless characterized by a neo-tectonic brittle dislocation system represented by the Aosta-Ranzola and Ospizio Sottile fault systems and by the emplacement along open fractures of ultrapotassic and calc-alkaline dikes (periadriatic magmatism related) and gold-bearing quartz veins (Ballèvre et al.1986; Bistacchi et al, 2001).

The complex structural-geological context of the Region deeply influenced the relief evolution and the slope dynamics. Superimposed to the tectonic evolution, the glacial morpho-dynamics - the region is the one with the largest glacierized area of Italy - have been influencing the slope setting, mainly due to the debuttressing caused by glaciers retreat (Giordan, 2018). Gravitational phenomena - shallow landslides, rockfalls, and large instabilities - played a fundamental role in shaping the geomorphology of the region. Rockfalls are highly frequent and represent a serious geo-environmental risk in the study area, crossed by the principal highway connecting Aosta to Piedmont Region in the south and to Switzerland, France, and the highly popular and frequented area of Valtournenche (Cervinia Ski Area) in the North. Due to its complex orography and extremely variable altitude range, the region is characterised by different climate regimes. According to the Köppen's climate classification (Figure 4.1d) (Rubel et al., 2017 and references therein), the climate of the area is classified as Alpine Frost (EF) and Alpine Tundra (ET) at higher elevations, as Boreal with warm summer (Dfb) in the central sector and on the slopes, as Warm Temperate (Cfb) along the secondary valley bottoms (mainly oriented N-S), and arid-semiarid (Bsk) in the eastern-central part of the Dora Baltea Valley bottom (E-W oriented).

Within the Aosta Valley region, the study area comprehends the Mountain Communities of Mont Cervin and Mont Emilius, which cover an area of 670 km^2 (bordered in Figure 4.1). The selection of this area was related to the availability of suitable rockfalls and meteorological datasets. The area covers an altitude range between 400 and 4500 m a.s.l., developing from the southern massifs to the northern peaks of the Region, and passing through the valley bottom nearby Aosta. Moreover, Mont Emilius main valley axis is E-W oriented, whereas Mont Cervin main valley axis is N-S oriented. Local factors such as orientation, inclination, location, land cover, snow cover and human infrastructures play an important role in the climatic and environmental local variability. The study area is characterized by the presence of different lithologies belonging prevalently to the ophiolitic Piemonte Zone and Austroalpine outliers (both lower and upper). In the southern part of the study area, metaophiolite-related lithologies (e.g. serpentinized peridotites, metagabbros and rodingitic dikes, metabasalts, metaquartzites, marble, and calcshists) of the Zermatt-Saas zone are abundant together with eclogitic terms belonging to the Mont Emilius Austroalpine outlier (e.g. metagabbros and metagranitoids and less abundant micashists). In the northern part of the Mont Emilius Mountain Community, the meta-ophiolite lithologies of the Combin Zone are prevalent (e.g. serpentinite/prasinite and flysch-type michashists), together with the Mont Mary Austroalpine upper outlier (e.g. granitoids and metagranitoids). Regarding the Mont Cervin Mountain Community, the valley bottom (i.e., Valtournenche) is prevalently composed by Zermatt-Saas and Combin zone lithologies, whereas the northern part of the valley is characterized exclusively by the Dent Blanche and Mont Cervin Austroalpine upper outliers (and others minoritarian fractions of the Pillonet

and Roisan outliers) comprehending prevalently metagranitoids, but also anphibolites and metabasalts. In the southern part of the Mont Cervin Mountain Community, the Austroalpine outlier of Chatillon-Saint Vincent outcrops; it is characterized by the same lithologies of the northern part, but with an eclogitic imprint.

4.3. Data

4.3.1. Rockfall dataset and its properties

The major limitation related to landslide inventories lays in the well-known inherent incompleteness of non-instrumental historical records of natural events (Guzzetti, 2000). The actual Aosta Valley landslide inventory "Catasto Dissesti Regionale SCT" (Figure 4.2a), publicly available at http://catastodissesti.partout.it/, represents the result of the integration of different sources, deriving from the evolution of the acquisition methods adopted.

During the 1990s, landslides information was recorded only qualitatively from different regional offices or scientific organizations. The great flood event of the Western Alps, occurred between 13 and 16 October 2000 (Ratto et al., 2003), represented a starting point for the regional authority to implement a more organized and structured landslide inventory, which comprised cartographic and quantitative information. During 2001-2004, Aosta Valley adhered to the construction of the Italian Landslide Inventory, named IFFI (Inventario dei Fenomeni Franosi in Italia, http://www.progettoiffi.isprambiente.it/cartografia-on-line/). From 2005, the regional database was enriched through field work and orthophotos observations. Even the historical phenomena and great paleo-events were further analysed and validated. Since 2010, the regional landslide inventory has been regularly updated thanks to an innovative automatic computerized data acquisition procedure, involving both the Regional Civil protection Department and the Forest Corps operating in the territory.

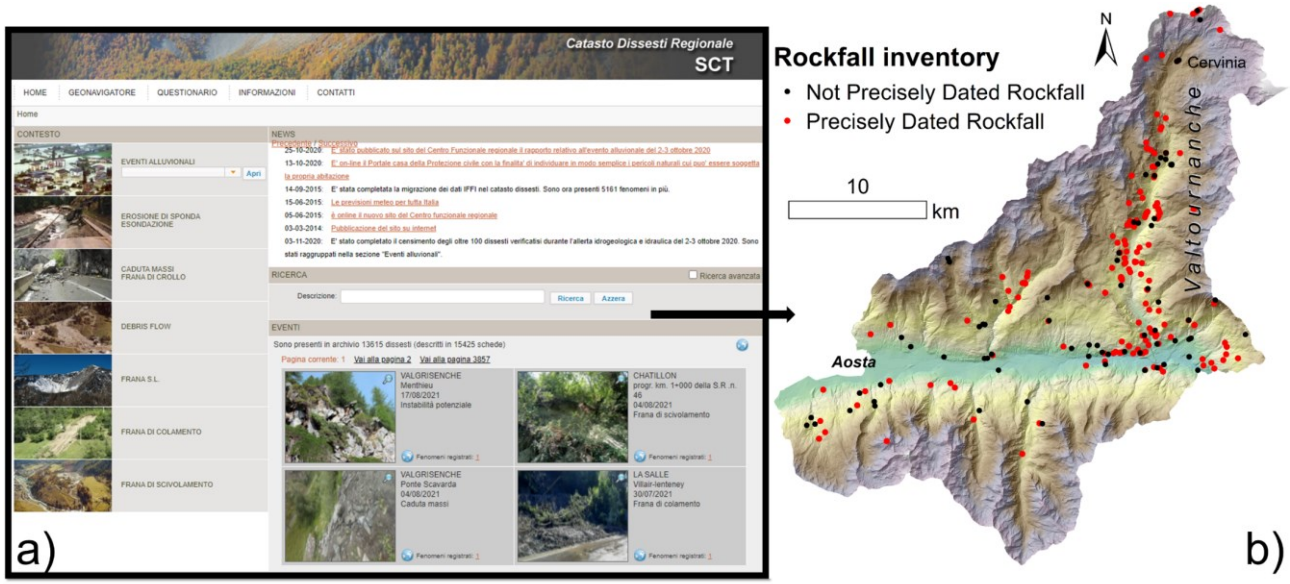


Figure 4.2 a) Visualization of the Aosta Valley region landslide inventory online portal. b) Available rockfall events locations in the selected study area.

Rockfall related data were extracted (Figure 4.2b) and thoroughly checked for errors, duplicated values, and for the identification of source areas when only deposit coordinates were recorded. In detail, each rockfall record was validated by consulting the Digital Terrain Model, orthophotos, geological and geomorphological maps available for the area (http://geoportale.regione.vda.it/), searching for source areas evidence (e.g., outcrops, scarps and rock walls). Once validated, each record was associated with a unique ID. For the reference period (1990-2018), 243 rockfall records were extracted from the Landslide Regional Database and validated. Almost 70% of the records (168 out of 243) come with the exact date of occurrence, whereas the remaining 30% (75 records) have only the month and year or only the year of occurrence. Histograms of rockfalls occurrence per year and per month were produced. The former was used to analyse the phenomena as a function of the acquisition method through time, the latter was expected to highlight seasonal patterns. Although the temporal distribution of the events in the study area (Figure 4.3a) suggested a slight increase in rockfall frequency in recent years (since 2000), such evidence could be linked to a reporting bias, following the establishment of more sophisticated monitoring networks and technologies. This issue was already pointed out by several authors working on alpine rockfalls (Nigrelli et al., 2018 and references therein) and inventory bias is a common and known problem when dealing with such natural hazards (Guzzetti et al., 2012; Petschko et al., 2016; Steger et al., 2016a). The seasonal characterization of rockfall occurrence (Figure 4.3b) showed an almost bimodal distribution of the events. The highest peak is recorded in spring (March, April and May) and a secondary peak is observed between October and January, with a slight decrease in December.

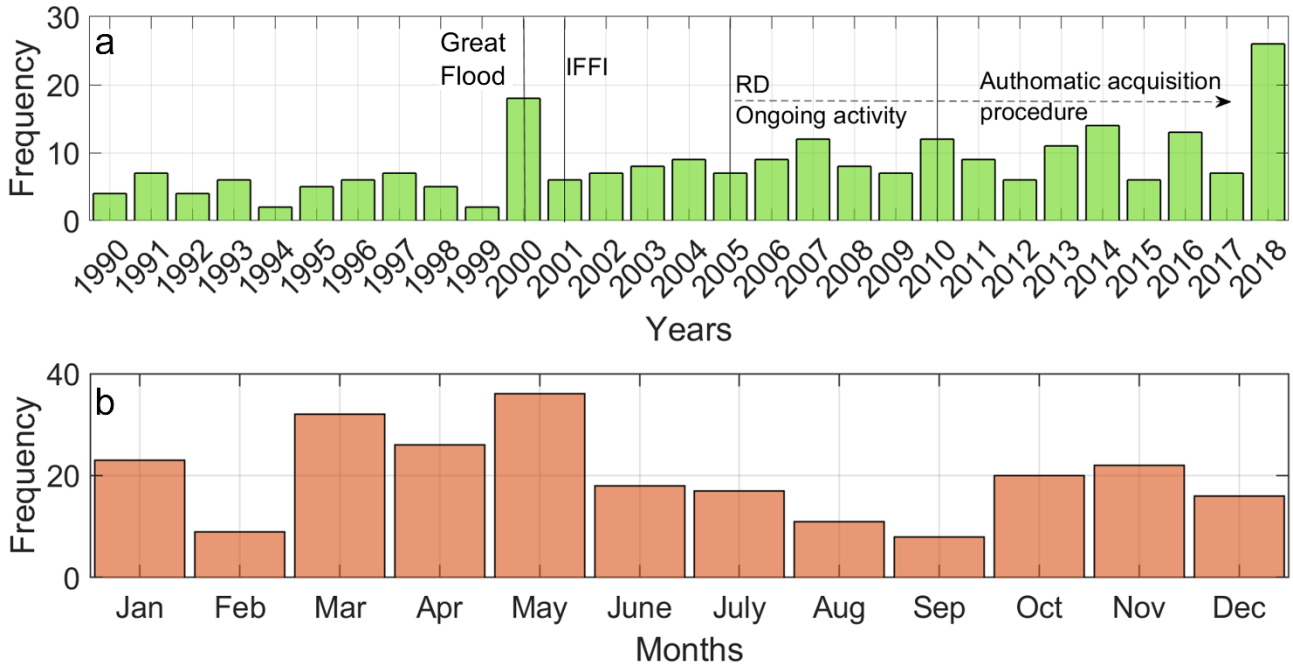


Figure 4.3 Annual frequency of rockfalls in the study area and historical evolution of the Regional Database (RD); b) Monthly frequency of rockfalls in the study area

A non-parametric analysis of the rockfall volumes was also carried out by means of histogram-based frequency and of the Empirical Cumulative Distribution Function (ECDF), which provided information on the probability of observing a rockfall of a certain volume. Only 199 rockfalls have an associated volume information, which in some cases is only a magnitude order and not a precise value. The most frequent volume class resulted to be 5-50 m^3 (Figure 4.4a); according to the data available in the inventory, the probability of occurrence of a rockfall with a volume larger than 10 m^3 is about 0.35, whereas the probability to observe a rockfall with a volume higher than 100 m^3 is of 0.15 (Figure 4.4b).

In terms of altitudinal distribution (Figure 4.4c), events appear to concentrate along the mid altitude range (400-2000 m a.sl.). This relative high abundance of rockfalls at these altitudes can be explained with the difficulties in monitoring high mountain environments, unless characterized by human activities (e.g., ski resorts, hiking itineraries). As a demonstration, in the area around Matterhorn - famous for hiking trails, climbing, and the ski resorts of Cervinia and Zermatt - dated events are available even at very high altitudes (>3200 m a.s.l.).

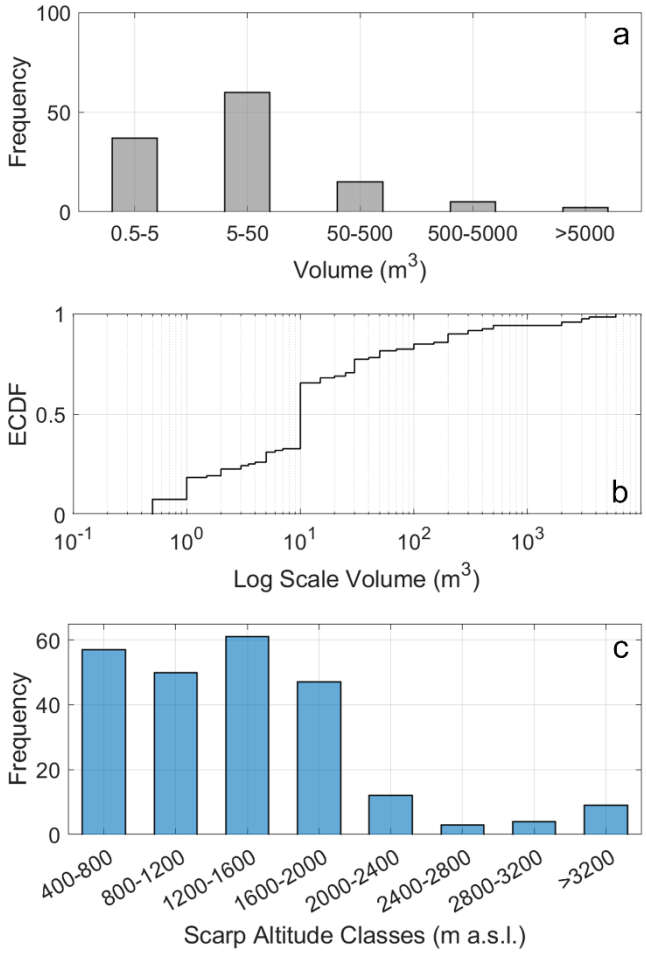


Figure 4.4 a) Frequency histogram of rockfall volume classes. b) Empirical Cumulative Distribution function (ECDF) of rockfall volumes. c) rockfall frequency vs scarp altitude

4.3.2. Land surface data

For this study the Digital Terrain Model (DTM) made available by Aosta Valley region (https://geoportale.regione.vda.it/mappe/informazionigeoscientifiche) was used. It has a 2 m x 2 m horizontal resolution and was resampled at 10 m × 10 m, using ESRI ArcGIS® 10.2.2 Spatial Analyst tools.

For geology, the 1:10000 geological-geomorphological map publicly consultable on the Aosta Valley geoportal WebGIS (http://geologiavda.partout.it/cartaGeologicaRegionale?l=it) was used; the associated shapefiles were made available upon request to the Regional Geological Office. The map reported different information levels, including the main geo-structural domains together with a more detailed description based on lithologic characteristics and metamorphic imprint. The map also reported quaternary deposit types, including "large boulder deposits", and linear geological-geomorphological features, including some landslide "detachment scarp". However, these deposits were not dated nor associable to the rockfall dated inventory described in Section 4.3.1 and the scarps were not classified depending on landslide type nor associated to the deposits with a univocal code; therefore, these data were not integrated with the rockfall dated inventory.

4.3.3. Climatic data

4.3.3.1 Weather station dataset

A total of 19 meteorological stations are located within the study area (Figure 4.5). They record temperature (T), precipitation (P) and snow height (Hs) data with different temporal coverage and resolution. The stations are located at altitudes from 400 to 3000 m a.s.l. and in different microclimatic domains (i.e., different slope, aspect and geomorphological-vegetational conditions). Starting in 2002, manual stations installed during the last century have been gradually dismissed and replaced by automatic ones (see also Table 4.1 for details). Meteorological data are publicly available on the website of the Agenzia Regionale per la Protezione Ambientale (ARPA, http://www.arpa.vda.it/it/) and on the website of the Centro Funzionale Regione Autonoma Valle d'Aosta (https://cf.regione.vda.it/portale_dati.php). For nearby stations, the automatically and manually collected data series were joined to obtain longer representative time series, following suggestions of ARPA technicians. Two auxiliary manual stations (ST14bis -Promiod Covalou and ST1bis-Ussin in Figure 4.5), considered homogeneous and complementary to the automatic stations ST14 and ST1 for

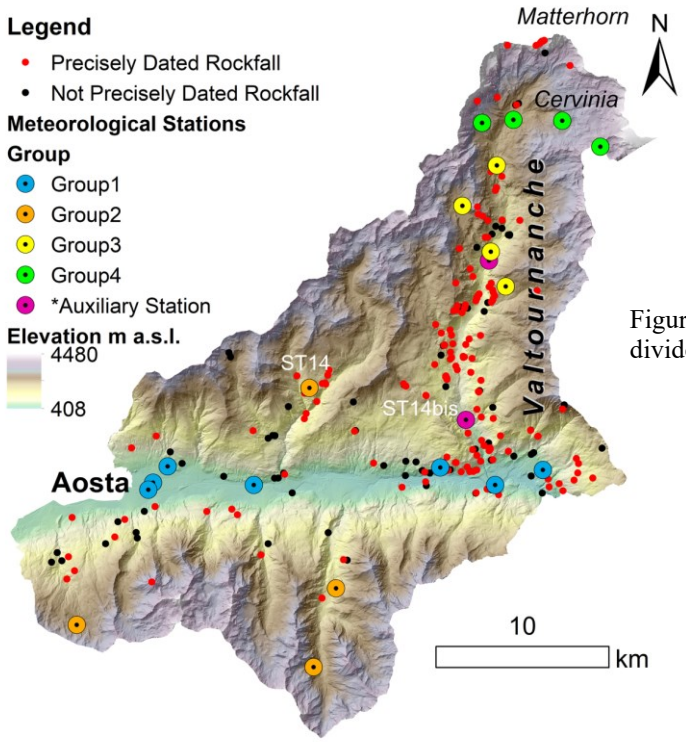


Figure 4.5 Available meteorological stations locations divided by group and rockfall locations.

precipitation by ARPA technicians, were also kept for rockfall-meteorological station association. After merging, a further manual quality check was carried out, to find out residual anomalous values and gaps (especially zeros instead of missing values) in the data series.

The oldest data series dates back to 1980 (ST17). Most of the manual stations were installed in 1990, whereas the automatic ones were installed in 2002. Manual stations present a daily temporal resolution, whereas the automatic ones collect data every 30 minutes. In some cases (namely ST2, ST13, ST16, ST7, ST9, ST10, ST11, ST12, ST28), not heated rain gauges were installed. Among these stations, those below 1500 m a.s.l. are characterized by missing data from December to March, while those located above 1500 m a.s.l. do not record from November to April. The available meteorological stations were split in four groups, depending on location (E-W lineament, N-S lineament) and altitude range.

- Group 1 – Stations located along the E-W lineament (Aosta Valley) at an altitude range between 400 and 850 m a.s.l. – ST2, ST3, ST4, ST13, ST16, ST17, ST28.
- Group 2 – Stations located along the southern slope of the E-W main valley (Aosta Valley) at an altitude range between 1500 and 2300 m a.s.l. – ST10, ST11, ST12. Station ST14 was also added to this group with a comparative purpose, since it is the only station located at a comparable altitude but on the opposite flank.
- Group 3 – Stations located along the N-S lineament (Valtournenche Valley) at an altitude range between 1300 and 2000 m a.s.l. – ST1, ST9, ST26, ST27.

- Group 4 – Station located at the head of Valtournenche Valley, at an altitude range between 2000 and 3100 m a.s.l. – ST5, ST6, ST7, ST8.

4.3.3.2 Precipitation and Temperature gridded dataset

The Centro Funzionale Regione Autonoma Valle d'Aosta developed a raster dataset (not published) of precipitation (mm) and Temperature (°C) for the whole Aosta Valley Region for the period 2003 to 2020. The dataset has an hourly temporal resolution and a spatial resolution of 0.00129° (about 122 m). Each map derived from the spatial interpolation of precipitation data from rain gauges (with no distinction between solid and liquid, thus including snow melting inputs for heated rain gauges) and temperature data from thermometers located on the regional territory. In particular, precipitation values were interpolated adopting the GRISO Rainfall generator (Pignone and Rebora, 2014). The dataset was made available upon request for this research in March 2021. The dataset came in the form of binary files. An example of how the not-processed gridded datasets appeared is provided in Figure 4.6.

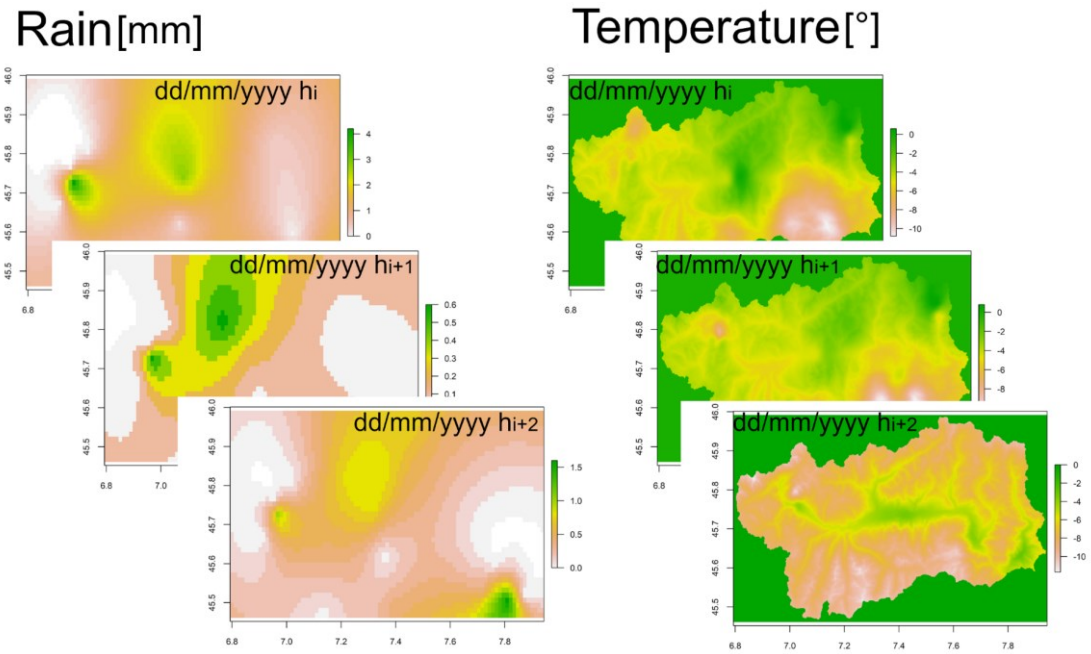


Figure 4.6 Schematic visualization of the available hourly grid-based precipitation and temperature datasets.

4.3.3.3 Snow Water Equivalent (SWE) gridded dataset

The Agenzia Regionale per la Protezione dell'Ambiente and the Centro Funzionale Regione Autonoma Valle d'Aosta developed a raster Snow Water Equivalent (SWE) dataset for the whole Aosta Valley Region for the period 2001 to present (Filippa et al., 2019). The dataset can be accessed on the ARPA website (https://www.arpa.vda.it/it/effetti-sul-territorio-dei-cambiamenti-climatici/neve/swe) and is made available in ascii format upon specific request. The dataset is limited to the winter months

(November-May) of each hydrological year, has a temporal resolution of eight days and a spatial resolution of 500 m. Each map represents, on a cell basis, the actual volume of water stored as snow expressed in terms of equivalent water height [m] as a result of the combination of the Snow Cover Area (SCA, derived from satellite data), interpolated snow height data (derived from station data, additional manual measures and topographic variables), and estimated snow density data (from manual measures). Due to the interpolated snow height factor, the dataset does not guarantee the conservation of mass between consecutive SWE maps (Camera et al., 2021). An example of how the not-processed SWE datasets appeared is provided in Figure 4.7.

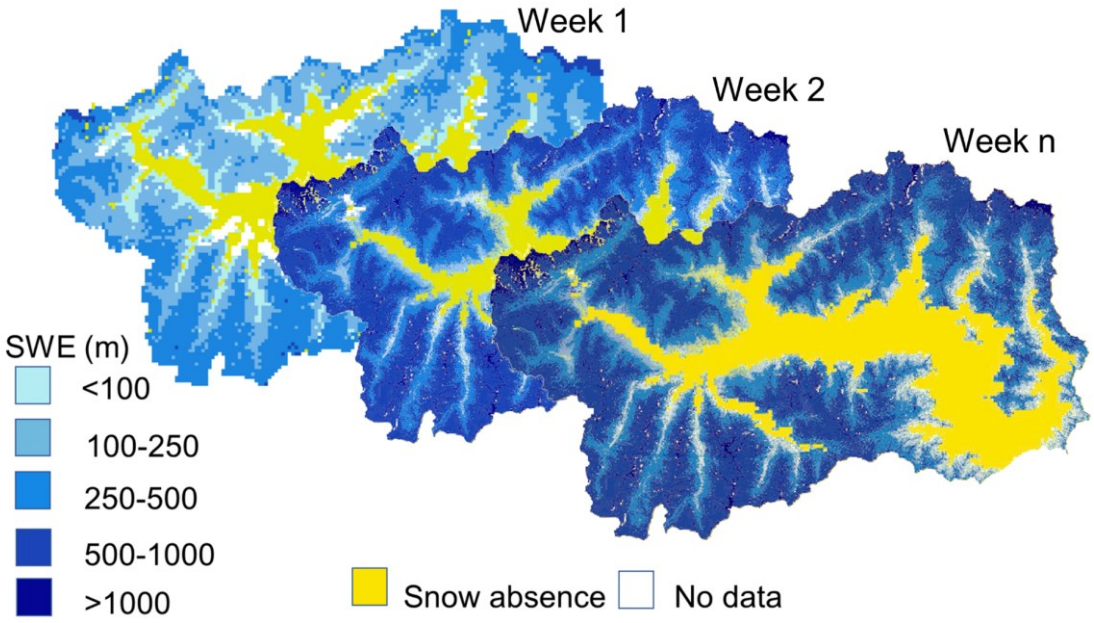


Figure 4.7 Schematic representation of the available grid-based SWE (Snow Water Equivalent) dataset.

4.4. Methodology in brief

The methodology followed three main steps (Figure 4.8). Accordingly, each step will be detailed presented in a separate Section followed by the associated results. The three-step procedure could be summarized as follows:

(i) Definition of a procedure to recognize climate conditions influencing rockfalls occurrence in an Alpine environment, combining both their critical short-term (triggering) and long-term (preparatory) effects. In detail, the study investigates the relationships between rockfall occurrences and four climatic indices, namely short-term rainfall (sub-daily data), effective water inputs (daily data, both considering rainfall and snow melting), wet-dry episodes and freeze-thaw cycles. This phase included the pre-processing of the station-based dataset, the calculation of the statistical distribution of the four rockfall related climate indices for the reference period, and the definition of empirical critical thresholds for each climate index. As the grid-based dataset presented in Section 4.3.3.2 was made available for this research in March 2021, it could not be used for this first step of exploratory analysis, conducted in 2019. This part of the research is presented in Section 4.5 "Deciphering meteorological influencing factors for alpine rockfalls: from station data" and published in the form of a scientific paper (Bajni et al., 2021a).

(ii) Calculation, for each index, of the mean annual exceedance frequency of the previous defined empirical critical threshold to be exploited as synthetic, spatially distributed climatic predictors in a rockfall susceptibility statistical model. These climatic predictors were produced both starting from a station-based hourly dataset, and consequent regionalization, and a grid-based hourly dataset. Additional snow dynamics related predictors were available from Camera et al. (2021) and introduced in the model. This part of the research is presented in Section 4.6 "Synthetic climatic predictors for rockfall susceptibility modelling".

(iii) Implementation of the climate related predictors in a rockfall susceptibility model, performed both using Generalized Additive Models (GAM) and investigation of their role, significance and technical usability in defining rockfall occurrence. In this regard, different issues regarding inventory bias, physical plausibility of climatic predictors and concurvity were addressed by stepwise modifications and improvements of the model setup, starting from a so-called "blind model" (i.e., a susceptibility model created without awareness of the rockfall inventory characteristics and flaws and of the physically driven processes potentially influencing susceptibility). This part of the research is presented in Section 4.7 "Rockfall susceptibility modelling".

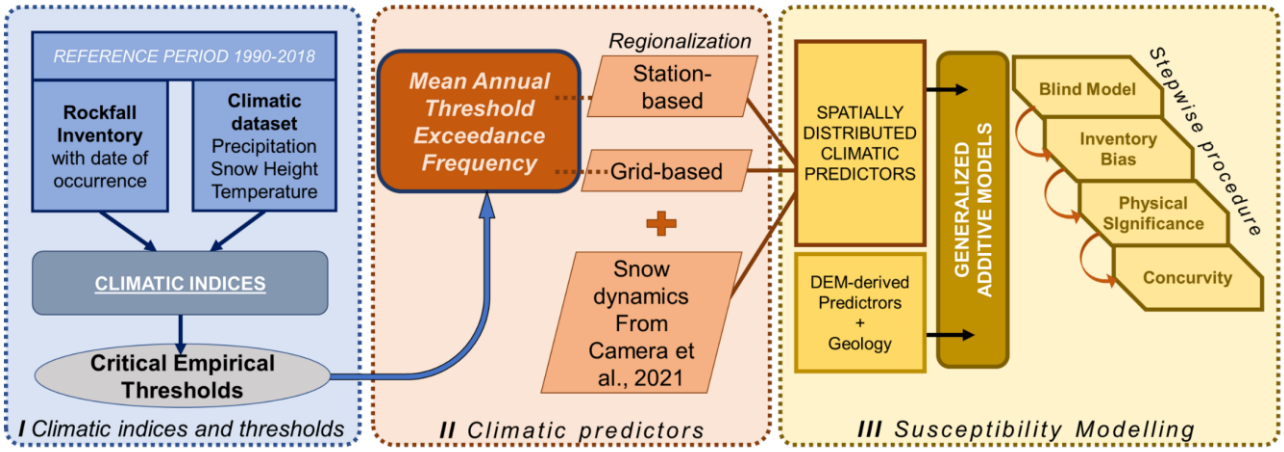


Figure 4.8 Synthetic workflow and procedural steps adopted for the Aosta Valley case study. Each coloured block represents a specific step and Section of the Chapter.

4.5. Deciphering meteorological influencing factors for alpine rockfalls: from station data processing to thresholds

4.5.1. Methods

4.5.1.1 Climatic data analysis and pre-processing: Altitudinal Temperature Lapse Rate (ATLR) and effective water inputs

A preliminary analysis of the daily aggregated T, P and Hs station derived data was carried out to verify climate differences and similarities between stations. Moreover, an altitude temperature lapse rate (ATLR) for the study area was evaluated on an annual base. As temperature data were available for all automatic stations only from 2010 to 2018, the ATLR was evaluated within this time period and considered valid for the entire reference period 1990-2018, following the concept of Climate Normals (WMO, 2007). A global ATLR for the study area was evaluated using mean annual daily temperature for the 2010-2018 period. Moreover, the covariable slope aspect was tested in addition to altitude to evaluate possible model improvements. Different ATLRs for the E-W lineament (Aosta valley) and for the N-S lineament (Valtorunanche Valley) were tested (using data from stations in groups 1-2 and groups 3-4, respectively) in comparison to the global ATLR. To compare the different models, the performance indices R-squared (R2), adjusted R-squared (Adj-R2, to compare models with a different number of covariables) and predicted R-squared (Pred-R2, to avoid model overfitting) were calculated. These comparisons were carried out only when a performance value (p-value) <0.05 for the model was verified. Different studies (e.g., Nishii et al., 2013; Crosta et al., 2014) pointed out that water supply deriving from snowfall can be quite relevant for the development of slope instability phenomena. A procedure to correct rainfall data in order to account for snowfall-deriving water inputs was thus implemented, with slightly different steps according to the station type. The different station types were identified based on temporal coverage, temporal resolution, instruments, and available data (Table 4.1).

The correction was carried out at the daily resolution, despite the availability of sub-hourly data (i.e. 30 min) for most stations. This choice was made because the daily time step already allows considering the occurrence of melting, freezing, and melting-refreezing processes - critical for snow melt contribution - by observing daily minimum and maximum temperature and their relationship with the 0 °C threshold. Further detail, such us including snow melting and refreezing processes at the hourly scale, was considered beyond the scope of this study. The correction procedure steps are described below.

Table 4.1 Station types, depending on sensors and data availability. T=Temperature; Phg=liquid precipitation derived from rainfall plus melted snow inside heated rain gauges; P=rainfall from standard rain gauges; Hs=snow height. ST9 belong to group B for data series between 1990 and 1998, and to group C for data series between 1998 and 2018.

Station Type	Sensors	Data	Station IDs
A	Heated Rain Gauge Snow Gauge Thermometer	Phg, Hs, T	ST1, ST5, ST8, ST26, ST27, ST17, ST14
B	Heated Rain Gauge Thermometer	Phg, T	ST3, ST4, ST9(1990-1998)
C	Rain Gauge Thermometer	P, T	ST13, ST9(1998-2018), ST16, ST10, ST28
D	Rain Gauge Thermometer Snow Gauge	P, T, Hs	ST7, ST11, ST12
E	Rain Gauge	P	ST2
F	Thermometer	T	ST6

Step 1: Identification of snowfall initiation representative temperature (*Ts*), degree day factor and average snow density from stations equipped with snow gauges and thermometers (i.e., type A and type D stations). For the definition of *Ts*, all positive temperature values corresponding to snowfall initiation were sampled (i.e., when the difference between *Hs* at day *i* and *Hs* at day *i-1* is positive, Hs_i-Hs_{i-1}>0 cm). A representative *Ts* was derived as the average of the sampled data. The calculated *Ts* value represents the temperature cut-off adopted to discriminate between solid (below *Ts*) and liquid (above *Ts*) precipitation in the subsequent steps. Following, daily *Hs* data were converted into Snow Water Equivalent (*SWE*) by using the classical equation (Seibert et al., 2015):

$$SWE = Hs\frac{\rho snow}{\rho water} \qquad \text{Eq. 4.2}$$

where *ρsnow* is snow density, and *ρwater* is water density. An average snow density value equal to 300 kg/m^3 was attributed based on literature about snow properties in the Italian Alps (Pistocchi, 2016; Guyennon et al., 2019). Water density was set at the commonly used 1000 kg/m^3 value. Following DeWalle and Rango (2008), a mean degree day factor was calculated for the study area. According to this approach, the daily melt rate *MR* (mm/day), expressed as equivalent water depths is given by:

$$MR = Cm * T_{mean} \qquad \text{Eq. 4.3}$$

where T_{mean} is the mean daily temperature and *Cm* (mm/degree-day C) is the degree day factor. At each station, to obtain daily *Cm* values, Eq. 4.3 was inverted setting *MR* equal to the previously calculated *SWE* (Eq. 4.2).

Step 2: Calculation of the actual daily liquid input (*Peff*) for type A and type B stations. Conceptually, the actual daily liquid input consists of the precipitation recorded at the heated rain gauge (*Phg*) deprived of the precipitation snow fraction (i.e., when $T_{mean}<Ts$) and with the addition of the snow melt from the accumulated snow (i.e., when $T_{mean}>0$ °C and $SWE>0$). Numerically, *Peff* was derived as follows:

$$Peff(i) = \begin{cases} 0, & Tmean \leq Ts \\ Phg(i), & Tmean > Ts \; and \; SWEr(i-1) = 0 \\ MR(i), & Tmean > 0 \; and \; Phg(i) + SWEr(i-1) > MR(i) \\ Phg(i) + SWEr(i-1), & Tmean > 0 \; and \; Phg(i) + SWEr(i-1) \leq MR(i) \end{cases} \quad \text{Eq. 4.4}$$

Where *SWEr* is the corrected *SWE* amount deprived of the melted fraction of snow, or with the addition of the precipitation snow fraction, and it is updated daily as follows:

$$SWEr(i) = SWEr(i-1) + Phg(i) - Peff(i) \quad \text{Eq. 4.5}$$

Step 3: Reconstruction of missing data (i.e., winter months data) for type C stations. It relayed on using a multiple regression procedure based on correlation of the type C stations available data with the closest Type A and Type B stations (i.e., one to three stations in a maximum radius of 9 km). A regression equation derived from the contemporary records was considered satisfactory for R-squared larger than 0.75. If this threshold was not reached, the type C station data were discarded. Once the time series were reconstructed, type C station data underwent the procedure described in Step 2.

Step 4: Conversion of snow gauge *Hs* data into *SWE* through Eq. 4.2 to obtain winter water inputs for Type D stations. This procedure was carried out only for days in which a melting episode was recorded (i.e., Hs_i-$Hs_{i-1}<0$).

Station ST2, is not equipped with a heated rain gauge (like group C stations) but was discarded from the procedure due to the completely lack of temperature data.

4.5.1.2 Climate indices definition and calculation

Using the original and processed time series, relationships between dated (at least date of occurrence) rockfall events and climatic conditions were explored. Rockfall events were assigned a reference weather station-based on two principles: (i) proximity, considered by using Thiessen polygons; and (ii) temporal coverage (i.e., meteorological data must be available on the day the rockfall event occurred and for the time windows relevant for the analysis). In case the two principles could not be mutually satisfied, the rockfall event was discarded from the analysis. Due to data availability (type, time coverage, time resolution), not all the four climate indices (STR, EWI, WD and FT, see Table 4.2 for definition) could be calculated at every station. Therefore, the configuration of the rockfall-station association and the number of stations and rockfalls included in the analysis resulted slightly different for each of the four

indices (Table 4.2). Figure 4.9 represents the three different configurations, obtained by applying the Thiessen (or Voronoi) polygon tool in ArcGIS®.

The main geomorphological limits corresponding to the crest dividing the two mountain communities, running from North to South in the middle of the map were maintained. For STR and WD stations two auxiliary stations were included in the configuration (ST14bis -Promiod Covalou and ST1bis-Ussin), which are manual stations considered homogeneous and complementary, in terms of precipitation, to the automatic stations ST14 and ST1 by ARPA technicians. They were used to fill gaps and to complete the oldest data series at the automatic stations. Moreover, they were used to densify the station network by simply projecting the same precipitation data series to the two auxiliary stations location. This was not reliably applied to the other indices, because they involved also temperatures and snow data, not available at the auxiliary stations.

Climatic conditions were defined in terms of four indices, which consider both short term (up to 120 hours) and long term (up to one year) periods. The four indices are: (i) STR, cumulated rainfall in 0.5-, 1-, 3-, 6-, 12-, 24-, 48-, 72-, 96-, 120-hours periods (effect of short term rainfall); (ii) EWI, daily, 3-, 7-, 15-, 30- and 60-days cumulated precipitation, including snow water inputs (effect of effective water inputs); (iii) WD, the number of wet and dry episodes in 30-, 60-, 120-, 180- and 365-days periods; (iv) FT, the number of freeze-thaw cycles in 1-, 3-, 7-, 15-, 30-, 60-, 120, 180- and 365-days periods. Time series of each index were calculated at suitable stations from 1990 to 2018 (or for the available time

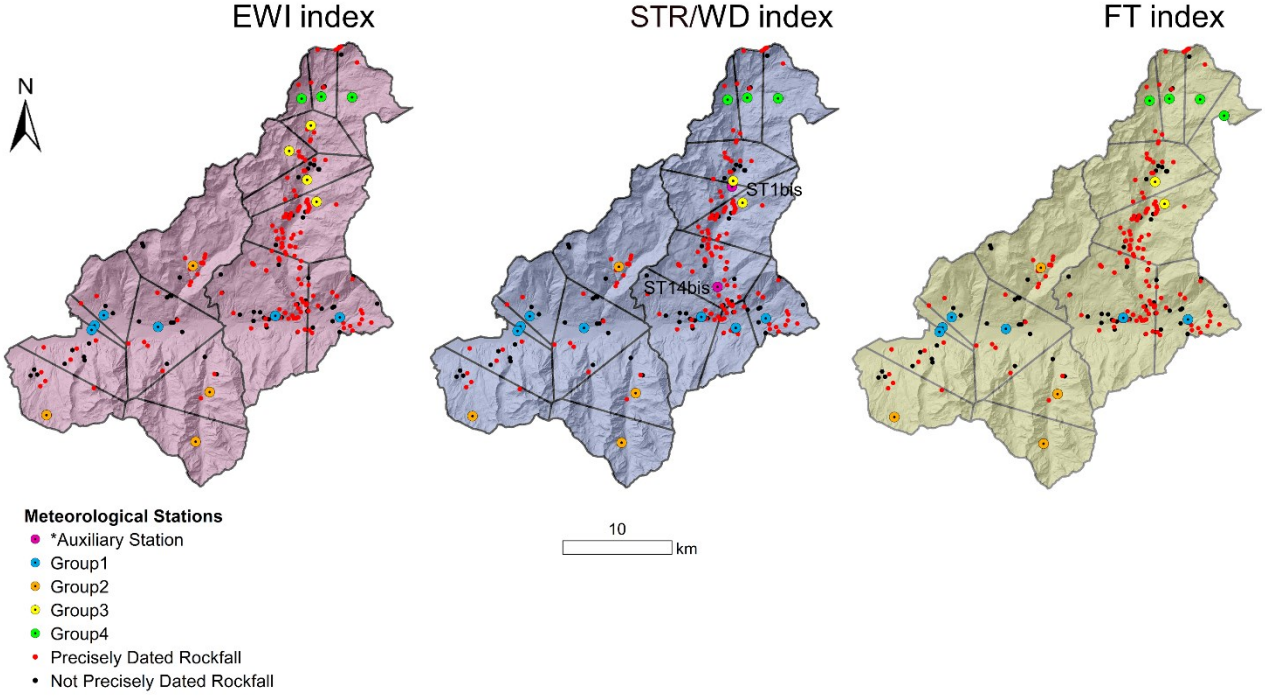


Figure 4.9 Associations rockfall-meteorological station for the calculation of the different indices done by Thiessen polygons.

period, if shorter) and summarized by means of statistical distributions. Not every station dataset covers the entire reference period 1990-2018, but based on the Climate Normals concept (WMO, 2007), the statistical distribution analysed for each station is considered representative of the entire reference period even when shorter.

For the calculation of STR and EWI indices, a rolling sum was applied for each chosen duration. For the STR index, the rainfall data series at 30-min resolution (i.e., both for heated and not heated rain gauges) were used as input, while for the EWI index the corrected daily Peff data series were employed.

Wet and dry episodes are calculated with a temporal resolution of 30 min. A wet episode occurs when rainfall in 30 min is ≥0.2 mm (that is the detection limit of the available rain gauges), showing at least an increasing humidity in the area. A dry episode occurs when in 24 consecutive hours no precipitation is detected. These criteria were set considering the available definition of independent rainfall events at the slope (macro) scale (D'Amato et al. 2016); albeit regarding intact rock specimens and not fractured rock masses, this interval is also included in the range of typical procedures at the laboratory (micro) scale, which includes several different drying intervals from 12 hours to 7 days depending on rock type and drying temperature (Kegang et al., 2016; Yang et al., 2019 and references therein). The WD index is expected to have an influence on rock weathering and mechanical weakening after a consistent number of cycles, even on the single laboratory sample; to be able to observe an influence at the rock wall scale it is reasonable to amplify the observation window in terms of time duration. Furthermore, to avoid a misleading superimposition of this index with the cumulated precipitation indices the smaller duration investigated was 30 days.

The FT index was based on T data with a temporal resolution of 30 min. A cycle is defined as a temperature transition across 0 °C, in both directions (from positive to negative and from negative to positive). Available data refer to air temperature and not to rock temperature, which is not available for the area and very difficult to acquire with good spatial and temporal resolutions (Nigrelli et al., 2018). However, for each rockfall, the temperature data series of each reference station were extrapolated to the failure zone altitude by applying the ATLR (see Section 4.5.1.1). Therefore, for each rockfall scarp location, the related statistical distribution of the index was produced applying ATLR to the data series of the associated station. Moreover, as for the wet and dry episodes, not ordinary conditions are expected to be more frequent in the longest time periods. Nevertheless, all the time windows used for the other indices were analysed to avoid neglecting the joint effect of freezing and thawing of water producing rock mass fatigue in shorter durations.

For WD and FT indices, a rolling sum technique was applied for each duration. To build their distributions, only records in correspondence of transitions in the number of cycles (episodes) were accounted, (i.e., consecutive repetitions of the same number were disregarded). This computational scheme was adopted to avoid multiple counting of the same cycle (episode).

The distribution for the STR index was obtained by calculating cumulated rainfall for each duration, then summarized by monthly maxima boxplots; this choice is justified to account for rainfall seasonality as short-term events are considered for this index. For the EWI index an Empirical Cumulative Distribution Function (ECDF) representation was preferred to summarize the statistical distribution related to each station, as it is a complete description of the sample (without a monthly distinction as durations involved transcend monthly subdivisions). For WD and FT indices a boxplot representation was preferred to summarize the statistical distribution related to each station, as it is more suitable for discrete data as cycles.

Table 4.2 Rockfalls-stations schematic configuration depending on the climate index

Index	ID	n. of rockfalls (out of 168)	Temporal resolution	Starting parameter	Selected durations	Station IDs
Short term rainfall	STR	96	30 min	P/Phg	0.5-, 1-, 3-, 6-, 12-, 24-, 48-, 72-, 96-, 120-hours	ST1, ST2, ST3, ST4, ST5, ST7, ST8, ST9, ST10, ST11, ST12, ST13, ST14, ST16, ST28
Effective water inputs	EWI	138	1 day	Peff	1-,3-, 7-, 15-, 30-, 60-days	ST1, ST3, ST4, ST5, ST7, ST8, ST9, ST10, ST11, ST12, ST13, ST14, ST16, ST28, ST26, ST27
Wet and dry episodes	WD	95	30 min	P/Phg	30-, 60-, 120, 180- and 365-days	ST1, ST2, ST3, ST4, ST5, ST7, ST8, ST9, ST10, ST11, ST12, ST13, ST14, ST16, ST28
Freeze-thaw cycles	FT	117	30 min	T	1-, 3-, 7-, 15-, 30-, 60-, 120, 180- and 365-days	ST1, ST3, ST4, ST5, ST6, ST7, ST8, ST9, ST10, ST11, ST12, ST13, ST14, ST16, ST28

4.5.1.3 Definition of not ordinary conditions

For each rockfall, the climate indices values associated with the triggering time of the event were extracted from the reference station time series. For the calculation of the STR index value at triggering

time, when the date but not the hour of the rockfall event occurrence was known, the entire day dataset was considered (i.e., until 12 pm). Then, the highest cumulated value of the correspondent day of occurrence of each duration was associated with the event.

As the goal was to investigate the severity of the triggering conditions, index values at triggering were evaluated in relation with the correspondent index distribution calculated at the reference station or, alternatively, in comparison to a global reference sample (see Section 4.5.1.4 for additional details). The comparison was carried out for each considered durations (as the critical duration is not a priori known) and it was evaluated at which percentile of the distribution the index value associated to the event triggering time corresponded. This procedure aimed at identifying if rockfalls occurred during not ordinary conditions (i.e., in correspondence of high percentiles of the distribution) definable as possible triggering conditions if the associated duration is short, or as possible preparatory conditions (in terms of repeated stresses perpetuation) if the associated duration is longer.

For each climate index, the critical percentile to recognize a condition as not ordinary was defined as the value exceeded by at least the 50% of the rockfalls and corresponding at least to the 75^{th} percentile but not over the 90^{th} percentile of its distribution. These criteria were chosen for two reasons. First, the procedure aims at accounting not only for extreme events but also for severe ones, potentially neglected with too high percentile cut-offs (e.g., 95^{th} or 99^{th}, usually adopted for extreme events analysis; Camera et al., 2017b). Second, considering 50% of the rockfall population for the recognition of the critical percentile should allow to include events occurred in severe (not extreme) conditions. Furthermore, if a consistent number of rockfall events (50% or more of those considered for each index) could be linked to specific climatic conditions, these conditions can be considered a reliable causation of the preparatory or triggering effect of climate on rockfalls.

4.5.1.4 From indices to climatic thresholds

For those indices respecting the criteria of percentage of events and percentiles presented above (Section 4.5.1.3), empirical thresholds (Intensity-Duration and number of cycles-duration thresholds) were defined. In case a specific rockfall presented index values higher than the critical percentile for more than one duration, the value closest to the critical percentile was considered and used for threshold construction. For the actual construction of the intensity-duration curve, the selected couples of values were interpolated using the non-linear least square fit algorithm of the Matlab® Curve Fitting Toolbox™ to adhere to a power law function. Following, shape and coefficient were adjusted by trial and error to define the most appropriate function so that the couples of values representing rockfall related to not ordinary conditions would fall above the resulting curves, except for some evident lower outliers.

These thresholds were constructed based on the previously defined critical percentiles exceedances in respect to the distribution associated to the single stations (i.e., "local approach"). However, when this type of Intensity-Duration (ID) thresholds is defined for a specific area, they cannot be easily exported to neighbouring regions due to meteorological and climate variability not included in the ID thresholds (Guzzetti, 2007). This is typically the case of mountain regions. Thus, to readjust data coming from different microclimatic conditions (e.g., valley bottoms and high peaks), two approaches were tested. First, a global approach was explored. Analogously to the local approach, the couples of values for threshold construction were selected by comparing rockfalls related indices values with the critical percentile value of the correspondent distribution, which in this case was built up using data coming from all the stations joined together. This global threshold was considered reliable only if similar to the local approach based one. If the global threshold was not considered reliable (i.e., too influenced by specific stations characteristics), normalized thresholds were analysed. For the precipitation related indices (STR, EWI) two types of normalization were carried out. The first threshold normalization was performed based on the Mean Annual Precipitation (MAP), while second normalization was based on the Rainy Day Normal (RDN), which is the ratio of the MAP to the mean annual number of rainy days (Guzzetti et al., 2007, 2008; Peruccacci et al., 2017; Leonarduzzi and Molnar, 2020). A normalized threshold was built also for the freeze-thaw cycles, taking inspiration particularly from the RDN normalization. It was defined as an intensity (of cycles)-duration threshold and introduced the Free-Thaw Normal (FTN) normalization factor. Analogously to the RDN, the new Freeze Thaw Normal (FTN) parameter is calculated as the ratio between the mean annual number of FT cycles and the mean annual number of across zero days (i.e., days where daily Tmax is positive and Tmin is negative). The threshold normalized to the FTN was thought to allow the comparison of rock walls at different elevations, which are exposed to different temperature regimes. Moreover, differently from normalizing only for the parameter mean annual freeze-thaw cycles (analogue to the MAP), it inherently accounts for whether the cycles are concentrated in some periods of the year, or if they are homogeneously distributed throughout the year.

4.5.1.5 Exploring the role of positive temperatures and gradients

Although of secondary importance in the alpine context (see Section 4.1), the role of positive temperatures in thermal stress induced fracturing and rockfall initiation, is worth to be investigated. An exploratory analysis regarding high temperatures was set up for those rockfalls resulting not associated to any of the previous described indices. Indeed, summertime rockfalls that seem unrelated to any climatic triggering-preparatory factor could be explained with not-ordinary positive temperatures and temperature gradients (Collins and Stock, 2016). In particular, following a similar approach as for the

other indices and according to findings of previous studies (Collins and Stock, 2016; Collins et al., 2018), the number of days, before rockfall occurrence, in which maximum daily temperature (Tmax) was recorded above the 99th percentile of its complete distribution and above the 90th percentile of its July-August distribution were calculated. Furthermore, the number of days, before rockfall occurrence, in which the daily temperature gradient (ΔT) was recorded above the 90th percentile of its distribution was calculated. As for the previous indices, distributions were calculated with all the available data for the reference period 1990-2018 and in fixed time periods from 1 day to 60 days.

4.5.1.6 Rockfall characteristics and relationship with climate indices

Besides thresholds construction, by linking rockfalls to climatic indices, an analysis was carried out to recognize seasonal patterns in the role played by the different triggering and preparatory factors. The relationships between volume, altitude classes and climate were investigated, too. To have a more intuitive comparison between the different classes (of both volume and altitude) in case of large differences in the number of samples (rockfall events), the analysis was carried out in relative terms, normalizing each term to the total number of rockfalls in the class.

Finally, possible relationships of rockfall abundancy and climatic indices associations with the underlying geological unit or lithology were investigated. Six geological units (refer to Figure 4.8c) were extracted from the Structural model of Italy at the 1:50 000 scale (Bigi et al., 1990) and from Ellero and Loprieno (2017), while seven rock types were extracted from the geological map at the 1:10 000 scale of the Aosta Valley Region (available upon request on the Aosta Valley Geoportal http://geologiavda.partout.it/cartaGeologicaRegionale?l=it). Rock types included (in order of areal abundancy): (i) Serpentinites-Prasinites; (ii) Meta-granitoids; (iii) Micashist and other metasedimentary Shists; (iv) Metabasites; (v) Calcshists; (vi) Marbles; (vii) Dolomites; (viii) Tectonized rocks; (ix) Sedimentary Breccias; (x) Quarzites.

4.5.2. Results

4.5.2.1 Altitudinal Temperature Lapse Rate (ATLR) and effective water inputs

The preliminary analysis confirmed that, despite general similarities between stations at comparable altitudes, there is a significant climatic variability within the study area. Local orographic-geomorphological dynamics seem important to determine differences in the rainfall and snowfall amounts, even between nearby stations. At high elevations, the northern head of Valtournenche valley (station Group 4 – see Section 4.3.3.1 and Figure 4.5) resulted to be particularly rainy (average annual precipitation 963 mm) with peaks mostly in May, August and November, and with snow intakes (average annual snowfall 694 cm) larger than those recorded on the southern slopes of the E-W lineament of

Aosta (station Group 2, average annual rainfall 774 mm with peaks in May and September-October-November and average annual snowfall 486 cm). Comparing the east and the west part of the study area, the former (Mont Cervin area, average annual rainfall 811 m) resulted to be more humid than the latter (Mont Emilius, area average annual rainfall 671 m). In general, at low-middle elevations along the E-W lineament (i.e., Group 1 stations), rainfall peaks are uniformly reached in Spring and Autumn (especially May and November).

Global and E-W Lineament ATLRs, evaluated with only the altitude covariable, resulted to have good model performance indices (Table 4.3). The global ATLR was −0.53°C/100 m and the same ATLR was obtained for the E-W lineament. Conversely, the ATLR for the N-S lineament resulted to be smaller (−0.43°C/100 m) but with a lower predictive performance (i.e. lower Pred-R^2, see Table 3). The slope aspect covariable resulted to added noise to the Global and E-W lineament ATLRs (i.e. lower Pred-R^2) but improved slightly the N-S lineament ATLR. The mean annual Global ATLR resulted to be the most reliable as it shows the highest regression predictive index (Pred-R^2) and it agrees with the values traditionally recorded in the Alpine region (Rolland, 2003).

Table 4.3 R-squared (R2), Adjusted R-squared (Adj-R2) and Predicted R-squared (Pred-R2) values for the tested annual ATLR

	Global		**E-W lineament**		**N-S lineament**	
Altitude	R^2	0.969	R^2	0.964	R^2	0.938
	Adj-R^2	0.967	Adj-R^2	0.959	Adj-R^2	0.922
	Pred-R^2	0.960	Pred-R^2	0.949	Pred-R^2	0.75
Altitude + Aspect	R^2	0.973	R^2	0.964	R^2	0.978
	Adj-R^2	0.969	Adj-R^2	0.953	Adj-R^2	0.963
	Pred-R^2	0.957	Pred-R^2	0.900	Pred-R^2	0.81

Regarding the effective precipitation calculation procedure, average *Ts* for single stations ranges between 1.8 °C and 2.8 °C, whereas the median *Ts* ranges from 1.4 °C to 2.4 °C. A *Ts* value equal to 2°C was thus chosen for the entire study area. For the degree day factor *Cm*, the mean value of the stations' medians (i.e. 3.5 mm/degree-day C) was selected as representative for the study area. Amongst the type C stations, only station ST9 was discarded in Step 3 (i.e. reconstruction of winter data series by means of multiple regression) because the implemented regression model did not reach the set criteria, with a R-squared equal to 0.66 (< 0.75). Due to these scarce results, no correction is thus carried out for data from 1998-2018. However, the station was not discarded for further analysis on the EWI index – rockfalls correlation. Firstly, ST9 almost unique position at the bottom of the middle Valtournenche Valley was considered crucial for the comparative analysis. Secondly, data from 1990 to 1998, coming from the old manual ST9 station, were suitable for snow melting inputs correction (i.e., Type B station).

Moreover, the majority of the rockfalls associated to ST9 from 2000 to 2018 (i.e., 6 out of 8) occurred in Summer and Autumn, thus without the strict necessity of snow melting inputs reconstruction. However, for this station the possibility of slightly underestimated ECDFs has to be considered in performing the consequent comparative analysis.

4.5.2.2 Short Term Rainfall (STR)

As a representative example Figure 4.10 compares, for station ST3, the maximum cumulated rainfall values in the different selected duration (i.e. 0.5 to 120 hours) before each rockfall event (if available, corresponding to the exact hour of occurrence, alternatively to the daily maximum) with the distribution of the monthly maxima cumulated values in the same duration and for the entire reference period. It shows that only 16 rockfall events (i.e., 17%) were associated to the exceedance of the 75th percentile of the distribution of at least one duration.

These results highlight that, except for some rockfalls associated to extreme events (such as heavy summer storms), in most cases short rainfall events alone are not able to justify the occurrence of rockfall events. Therefore, in the study area and for the considered volume range (more than 5 m^3), a direct cause-

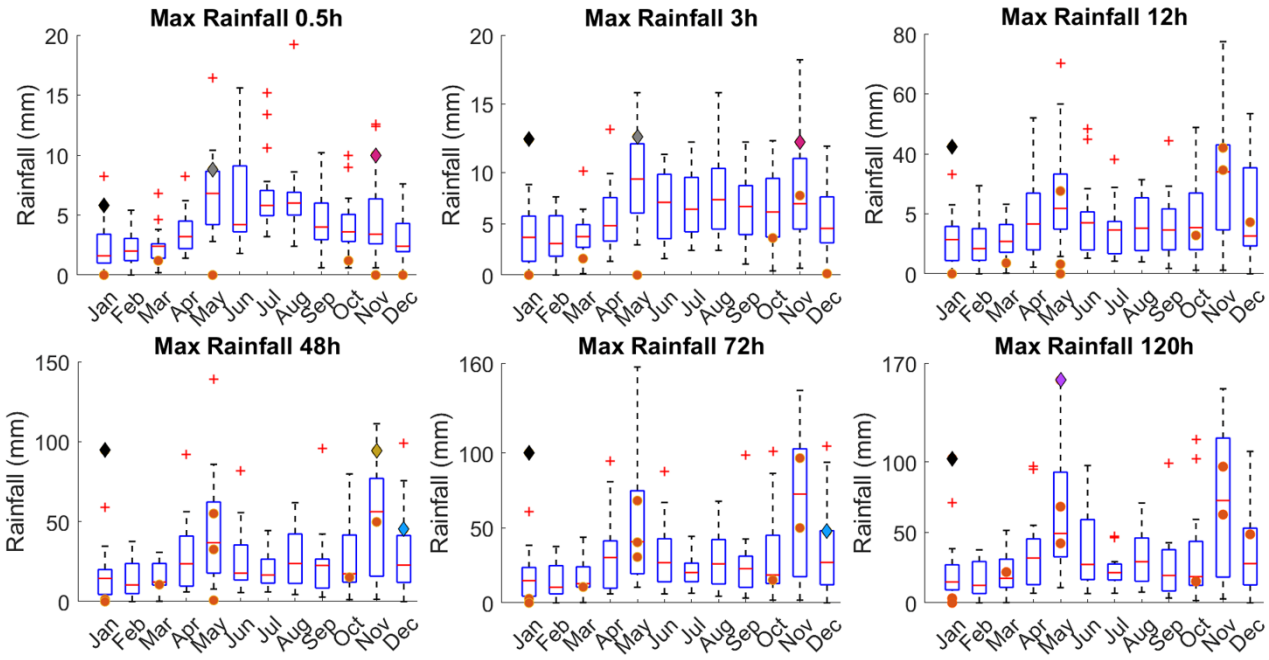


Figure 4.10 Boxplots of monthly maxima cumulated rainfall for selected duration from 0.5 to 120 hours. Points represent cumulated rainfall values corresponding to each rockfall event not exceeding the 75th percentile while diamonds represent cumulated rainfall values corresponding to each rockfall event exceeding the 75th percentile (each colour represents a specific rockfall and is maintained between the boxplots). All the presented graphs refer to ST3 data and associated rockfalls as an explicative example. The bottom and top of each box are the 25th and 75th percentiles, the line in the middle of each box is the median, outliers (crosses) are values whose distance from the box is higher than 1.5 times the interquartile range. Whiskers go from the end of the interquartile range to the furthest observation within 1.5 times the interquartile range.

effect relationship between rockfalls and short rainfall events cannot be assumed. This index, as it did not respect the procedure requirements adopted, was discarded.

4.5.2.3 Effective Water Inputs (EWI)

As a representative example, Figure 4.11 shows the ECDFs ensemble for station ST14, representing the distribution of the EWI index for the different chosen durations (i.e., 1 to 60 days) and the associated rockfall values. Despite the short-term rainfall index was discarded, by introducing durations of 1, 3, and 7 days in the EWI index, short and intense events were considered in the threshold construction, too. Carrying out the comparison with the ECDF curve, it resulted that more than 50% of the rockfalls were associated to at least one value larger than that corresponding to the 0.9 probability of the ECDF curve. This defined the 90th percentile as that representing not ordinary conditions for the area. More precisely, 73 rockfalls out of 138 were associated to not ordinary conditions. By considering the 75th percentile, the exceedance would increase to 80 rockfalls, leading to an increase in the relative number of rockfalls characterized by not ordinary conditions from 52.9% to 58.0%. However, the 90th percentile satisfied stricter criteria and it was therefore preferred.

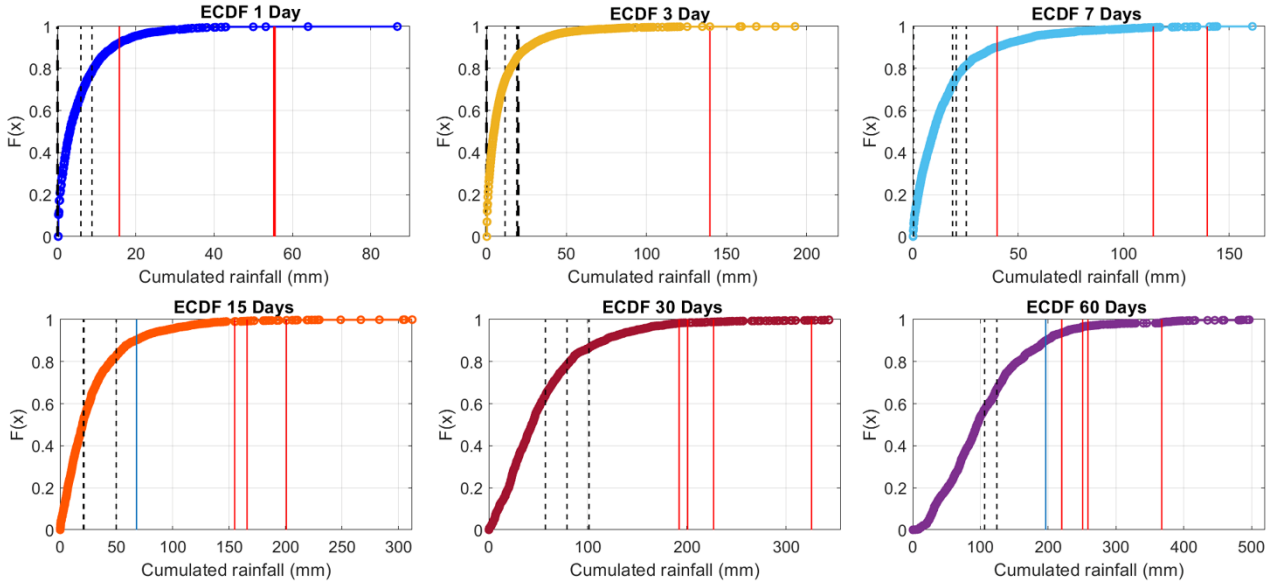


Figure 4.11 ECDF curves of cumulated precipitation for duration from 1 to 60 days. Red lines represent rockfalls with cumulated rainfall values over the 0.9 probability for more than one duration; blue lines represent rockfalls with cumulated rainfall values over the 0.9 probability for a single duration; dotted black lines represent rockfalls with cumulated rainfall values never exceeding the 0.9 probability value. The presented graphs refer to ST14 data and associated rockfalls.

The global approach was evaluated, too. Consistently with the local station approach, it showed more than 50% of the rockfalls associated to a value over 0.9 of the global ECDF curves, confirming the robustness of the adopted method. However, the global 0.9 cumulated rainfall values resulted to be very similar to those obtained for the stations with the longer temporal coverages, mostly located in the Mont

Cervin subarea (7 out of 9 stations). In particular, the adoption of the global approach, resulted in having (+3) events in not ordinary conditions in the northern part of Valtournenche valley, and (+5) in ordinary condition in its lower part. These resulting unbalanced relationships indicated a higher influence of data series of some stations in comparison to others in defining not ordinary conditions, and consequently in threshold definition. This led to prefer the local approach and consequent threshold normalization (see Section 4.5.1.4). Therefore, MAP and RDN normalized thresholds based on the local approach were calculated and evaluated in comparison to the not-normalized one (Figure 4.12). All the three thresholds are characterized by a negative power law. Figure 4.12 also shows that for very short (1 day) and very long durations (≥ 30 days) the not normalized threshold has a lower capability than the normalized thresholds to discern between ordinary and not-ordinary conditions. Moreover, over the threshold (Figure 4.12a), it is possible to distinguish the overlap of couples of I-D values characterized by percentiles either larger or lower than 0.9. This means that the normalized thresholds are more reliable than the not-normalized one since they smooth better microclimatic differences.

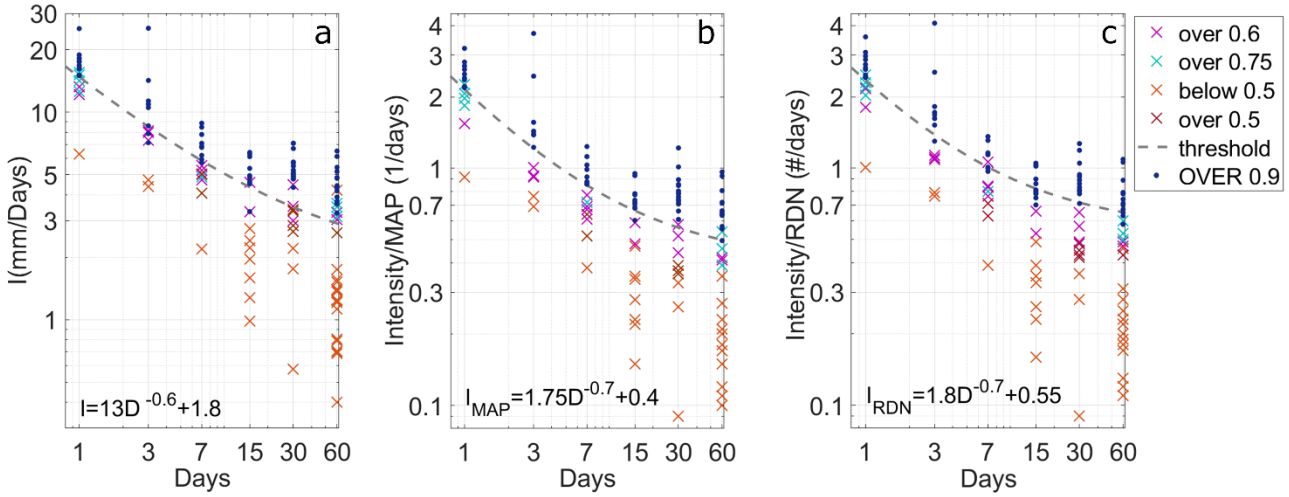


Figure 4.12 Intensity-Duration thresholds for the effective water input index. a) Standard threshold. b) Mean Annual Precipitation (MAP)-normalized threshold. c) Rainy Days Normal (RDN)-normalized threshold.

4.5.2.4 Wet and Dry Episodes

Wet and dry episodes were calculated with data covering the period 2002-2018 at almost every station and a temporal resolution of 30 min. No correction with snowfall data was adopted for this index, therefore a bias due to not heated rain gauges (9 out 15 stations used for the index) has to be considered for further applications beyond threshold construction. As a representative example, Figure 4.13 shows the boxplot ensemble related to the wet and dry episodes distribution (durations from 30 to 365 days) with the associated rockfall occurrence values, for station ST1. Carrying out this comparison for each station and associated rockfalls, it resulted that more than 50% of the rockfall (69 out of 95, 72%) were

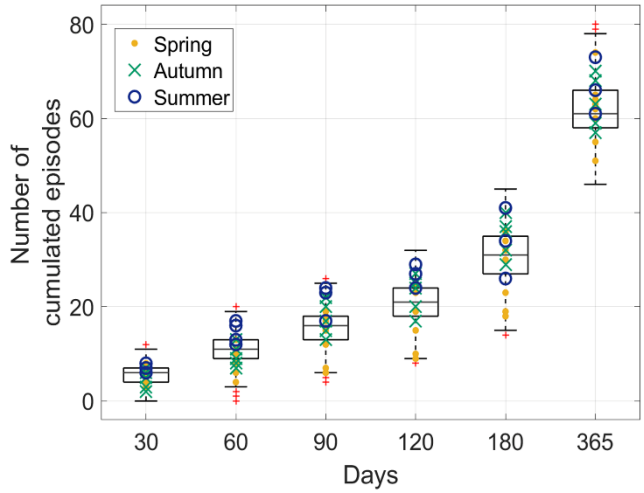


Figure 4.13 Boxplots representing the statistical distribution of wet and dry episodes for duration from 30 to 365 days. Superimposed symbols represent rockfalls associated values, divided by season of occurrence. Figure refers to ST1 data and associated rockfalls. Boxplot statistical limits as in Figure 4.10.

associated to at least one value over the 75th percentile of the distribution, thus defining not ordinary conditions for the area. By considering the 80th percentile the exceedance would decrease to 44% (42 rockfalls), while considering the 90th percentile the exceedance would decrease to 17% (16 rockfalls). The thresholds construction was thus carried out. Conversely to the EWI index, the global approach seemed applicable as the two thresholds resulted to be very similar (Figure 4.14). Moreover the 75th global percentiles resulted to be very close to the median value of the 75th stations associated values distribution (boxplots in Figure 4.14). The global threshold is almost correspondent to the threshold built with the classical station by station approach. This does not state that the number of wet and dry episodes is the same within the area, but only that number of episodes related to not ordinary conditions leading to rockfall phenomena are quite similar for the rock masses in the area. As a global threshold for the whole area was individuated for this index, no further normalizations were carried out. The threshold curves have the form of a positive power law.

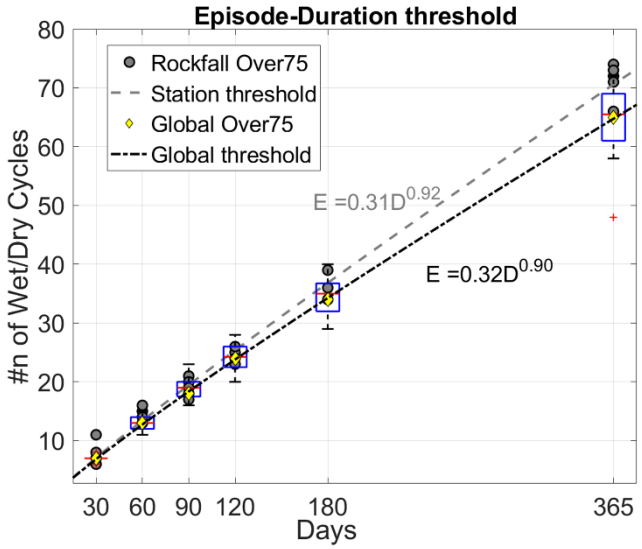


Figure 4.14 Episode-Duration Threshold for number of cumulated Wet-Dry episodes (30 to 365 days). Comparison between Global and local (Station-by-Station) approach. Boxplots represent the 75th percentile distribution among the stations. Boxplot statistical limits as in Figure 4.10.

4.5.2.5 Freeze-Thaw Cycles

Figure 4.15 shows, as an example, the boxplots ensemble of the freeze-thaw cycles distribution (duration from 30 to 365 days) and the associated rockfall occurrence values, for station ST8. Carrying out this comparison for each station and associated rockfalls, it resulted that more than 50% (64 out of 117, 55%) of the rockfalls were associated to at least one value over the 75th percentile of the distribution, thus defining not ordinary conditions for the area. Considering the 80th percentile the exceedance would decrease to 47.8% (56 rockfalls), while selecting the 90th percentile the exceedance would drop to 43.5% (51 rockfalls).

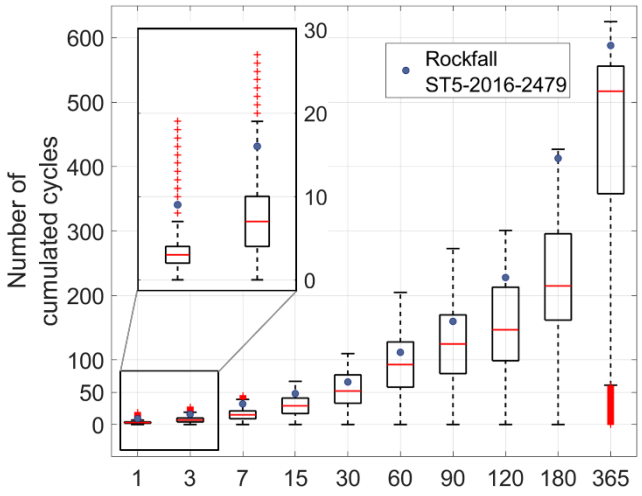


Figure 4.15 Boxplots representing the statistical distribution of freeze-thaw cycles for variable durations. Black dots represent the rockfall associated number of cycles. Example from rockfall ST8-2016-2479 (i.e. rockfall related to station ST8, occurred in 2016 with a scarp located at 2479 m a.s.l.). Boxplot statistical limits as in Figure 4.10.

Figure 4.16 displays the two thresholds built for the FT index. The first relates number of FT cycles (NoC) and duration and it was built from the original station data (Figure 4.16a); it has the form of a

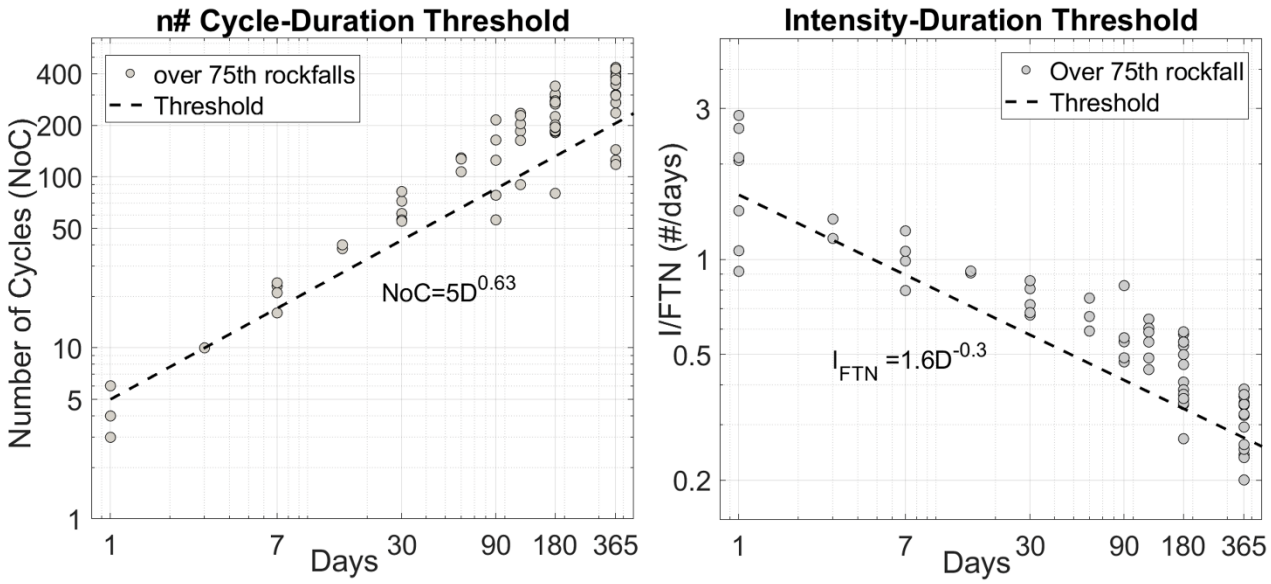


Figure 4.16 FT Cycles-Duration thresholds. a) Standard threshold, Log scale. b) FTN-normalized threshold Log scale

positive power law. The second threshold relates the intensity of the FT cycles, normalized to the FTN parameter (I_{FTN}), to the duration (Figure 4.16b). Coherently with the use of intensities instead of number of cycles, the resulting curve has the form of a negative power law.

A global threshold in this case was not considered because it would have given much more weight to the highest elevations, hiding not ordinary conditions exceedance for low-mid altitude rock walls. Temperature, differently from rainfall, is a continuous parameter and rockfalls are an expression of the loss of equilibrium between rock masses and the external environment, including the oscillation frequency through 0°C. For these reasons, the most suitable approach seemed to be a normalization, which should highlight anomalous oscillation frequencies for specific local conditions. However, even the normalization was not able to capture entirely the extreme local variability of the phenomenon of freeze-thaw cycles. As shown in Figure 4.16 some rockfalls are evidently very distant from the threshold curve, both up and down; this means that the threshold could both underestimate and overestimate the process. A slight improvement with the normalized threshold could be observed for middle-long durations (i.e., 120 and 90 days), but a high dispersion persisted for the shortest and longest durations (i.e. 1 and 365 days).

4.5.2.6 Deciphering triggering and preparatory climatic factors

Out of the 168 dated events, 136 could be linked to climate indices. The remaining 32 rockfalls were discarded due to the lack of complete climate data (e.g., missing data in the time window related to the index calculation). Among these 136 rockfall, more than 95% (130) resulted to be associated to not

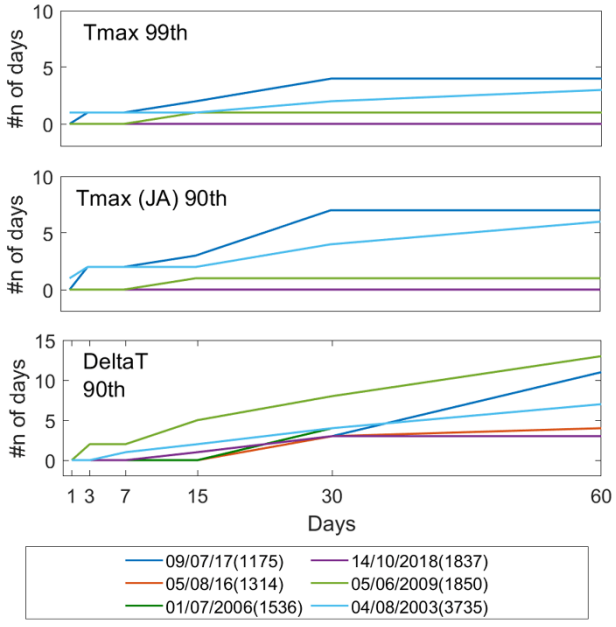


Figure 4.17 Number of days before rockfall occurrence, in durations from 1 to 60 days, in which Tmax were above the 99th percentile (*Tmax 99th* if the distribution was built with all the daily Tmax values) or the 90th percentile (*Tmax (JA) 90th* if only based on July and August) and daily gradients was above the 90th percentile (*DeltaT 90th*).

ordinary conditions of effective water inputs, wet and dry episodes, freeze-thaw cycles or to a combination of these factors. Only six rockfalls occurred during ordinary meteorological conditions for EWI, WD and FT indices. However, five of these occurred during Summer and one in early Autumn, so their possible relation with positive temperature anomalies was analysed (Figure 4.17).

In particular, a rockfall that occurred in July 2017 at 1175 m a.sl. and a second one that initiated in August 2003 at 3735 m a.s.l. showed a high number of days, in different durations before their occurrence, in which daily Tmax was above the 99th percentile of the complete Tmax distribution or above the 90th percentile of the July-August Tmax distribution. Also, they showed a quite large number of days in which the daily ΔTs were over the 90th percentile of their distribution, together with another rockfall occurred on 05/06/2006 at 1850 m a.s.l. For the other three rockfalls - which took place at medium altitudes on 05/08/2016, 01/07/2006 and 14/10/2018, respectively - an influence of high temperatures and gradients was more questionable, even if they showed an increment of the number of days in which the daily gradient was above the 90th percentile in periods of 30 days and 60 days before their occurrence.

Figure 4.18 analyses the seasonality of the triggering and preparatory factors. It shows that:

(i) During the Spring season, the most frequent factor is made up of a combination of precipitation characteristics (EWI and WD indices) and FT cycles, followed by FT alone and EWI (or EWI combined with WD episodes). In Spring, FT cycles could be linked, alone or in combination with rainfall, to 70% of the rockfall, stressing their predominant role in influencing the occurrence of such phenomena.

(ii) During Summer the combination of a precipitation related index and FT cycles is still the most frequent association of factors leading to failure, reasonably due to the lag between stabilization of T above 0°C and snow meltwater inputs still active at high altitudes. However, also WD cycles and effective precipitation have a high frequency. Some rockfalls could also be associated with high temperature peaks.

(iii) During Autumn precipitation is reasonably the main triggering-preparatory factor for rockfalls in the area, in correspondence with the second annual peak of the bimodal distribution of rainfall.

(iv) During winter months, a prevalent factor is not easily recognizable. However, effective precipitation seems crucial, especially in terms of snowfall income subject to intermittent accumulation and melting phases, or in combination with FT cycles, thus with freezing, melting and refreezing process.

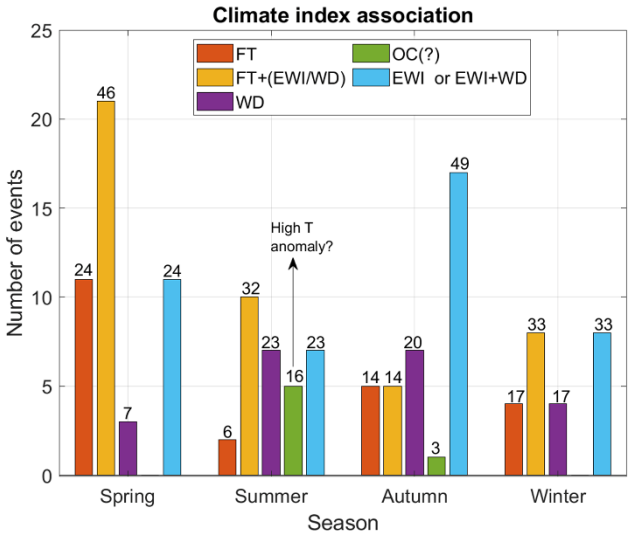


Figure 4.18 Bar chart representing triggering and preparatory indices seasonal frequency. FT= Freeze Thaw Cycles; EWI=Effective water inputs; WD=Wet and Dry cycles; OC=Ordinary Conditions. Numbers refer to the relative percentage of events of each season.

This analysis also highlighted that the choice of the 75th percentile minimum cut-off to consider not ordinary conditions occurrence resulted suitable, as it was reached and exceeded throughout the year with different frequencies according to seasonality. As considered durations include both short term (i.e., 1,3,7 days) and medium to long term (i.e., from 15 to 365 days) time periods, the seasonality of climate related rockfalls can be biased depending on the duration considered critical. Nevertheless, this is consistent with the attempt to consider climate not only as a trigger but also as a long-term preparatory factor, going beyond the single season.

Relating rockfall scarps altitude and associated climate factors (Figure 4.19a), a decreasing influence of EWI index could be observed by increasing altitude. Cumulated precipitation seems to be a very frequent triggering-preparatory factor at low-medium altitudes (400-1300 m a.s.l.). Conversely, WD episodes show a slightly growing effect on rockfall failure with increasing altitude. This could be linked to the different nature of precipitation events at high altitudes, where short and local storms prevail over high amounts of water resulting from long events. The association of FT cycles with EWI and WD episodes seems to remain stable with altitude. Finally, the single effect of FT cycles shows a peak at medium-high altitudes (1300-1800 m a.s.l.); this could be related to the transition zone where the 0°C temperature line is oscillating the most during the year.

Regarding the relationship between volumes and climate factors, only 73 out of the 136 analysed rockfalls have volume classes available, therefore only a partial analysis could be carried out (Figure 4.19b). The WD episodes effect is most evident for the smallest rockfalls (0.5-50 m^3). FT cycles alone and in association with precipitation related indices are predominant for rockfalls of medium and high

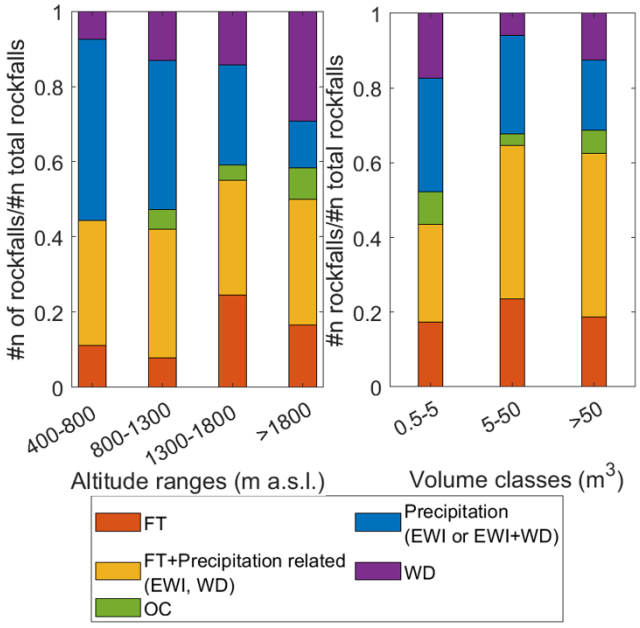


Figure 4.19 Bar charts representing triggering and preparatory indices and their relationship with (a) altitude and (b) volume classes. Data are normalized to the total number of rockfalls belonging to each class to make classes comparable.

volumes (larger than 5 m^3). EWI resulted to have a quite high effect both on small and medium size rockfalls. However, more data on volumes are needed to reliably investigate these relationships.

By comparing rockfall abundancy with geological units (Figure 4.20a), it resulted that the two most abundant units, the oceanic Zermatt-Saas and Combin units (38% and 33.8% of the study area respectively), gather the majority of rockfalls of the study area. This evidence is nonetheless confirmed in absolute terms of numbers of events per unit area (dots in Figure 4.20a). Considering climatic factors (Figure 4.20b), in the Zermatt-Saas unit the most abundant triggering-predisposing factors were represented by rainfall (EWI or EWI+WD), followed by FT+EWI/WD. In the Combin unit, the opposite occurs. Only 10 rockfalls were recorded in the C-MM-P unit, covering 17.3% of the study area, and for this reason an assumption on the prevalent climatic indices is hardly feasible; however, 4 out of 10 rockfalls are ascribable to WD alone. The other three units are less abundant and with none or just 1 rockfall recorded.

Considering lithologies rather than units (Figure 4.20c), it resulted that the number of rockfalls in Granitoids is much less than in Serpentinites-Prasinites, having these two lithologies almost the same abundancy (24.7% and 23.3% respectively) in the study area. In terms of number of events per unit area, Metabasites and Calcshists showed the highest value. Analysing the role played by climate (Figure 4.20d), for Serpentinites-Prasinites lithology, the most common influencing factor was represented by the association FT+EWI/WD, followed by EWI or EWI/WD. Calcshists lithology was mostly associated to rainfall (EWI or EWI/WD), which doubled the FT+EWI/WD combination: For Metabasites EWI or EWI/WD and FT+EWI/WD had almost the same abundancy. FT alone had almost the same abundancy

amongst the three above mentioned lithologies (slightly lower in Serpentinites-Prasinites). Although in Granitoids and Shists rockfalls are few and strong conclusions cannot be drawn, the events showed predominant relationships with WD and FT+EWI/WD, respectively. The association between climatic indices and underlying geological units and lithologies should be linked also to local rock mass properties such as in-situ stress and fracture network characteristics. Such an analysis would allow much more robust generalizations and the extrapolation of clearer relationships.

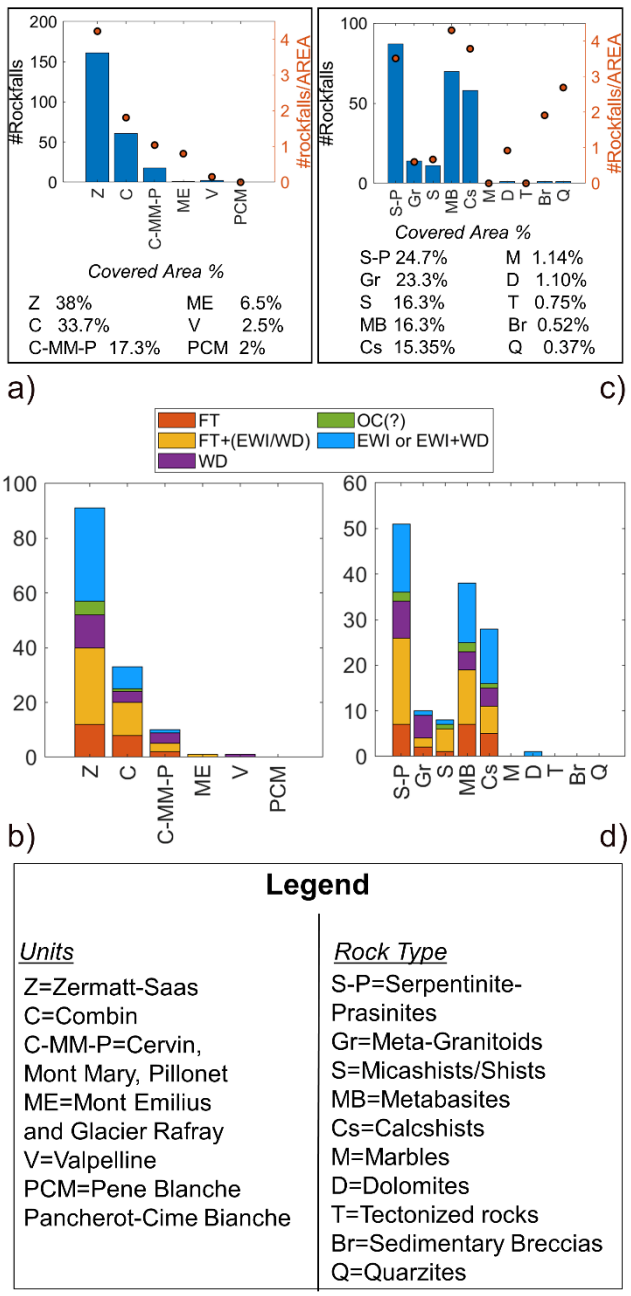


Figure 4.20 Frequency of rockfalls (both precisely and not precisely dated) in relationship with a) geological-structural underlying unit and c) lithology. Dots represent the frequency of rockfalls divided by the percentage of the area covered by the unit (lithology). Bar charts representing triggering and preparatory indices and their relationship with (b) geological unit (d) lithology/rock type.

4.6. Synthetic climatic predictors for rockfall susceptibility modelling

4.6.1. Methods

4.6.1.1 Calculation of climatic threshold exceedance frequency

Following the definition of the climatic indices, not-ordinary conditions and associated critical thresholds, it was necessary to translate these concepts into climate-related summary variables, suitable for the implementation of a statistical susceptibility model. Literature suggested two main approaches. The first approach was the one proposed by Gassner et al. (2015) and Knevels et al. (2020), who used multiple-day maximum cumulated precipitation as climatic predictors in their event-based landslide susceptibility analysis. The second approach, proposed by Catani et al. (2013), was based upon testing variables expressing the return period of rainstorms of given typologies, i.e., for which the critical rainfall amount or duration was a priori known or given by the Regional Civil Protection Office. Both approaches were not optimal for the present study, dealing with a multi-temporal analysis and for which prior knowledge was missing regarding specific critical amounts or durations for the previously recognized rockfall-related climate processes. Moreover, following the definition of not-ordinary conditions given in Section 4.5.1.3, the procedure is aimed at identifying climate processes either as triggering or preparatory conditions, based on the associated duration, or even as a combination of these two aspects at different temporal scales. Therefore, to effectively summarize the role of the climatic processes individuated with the procedure in Section 4.5, an approach similar to Camera et al. (2021) was implemented. For each index, starting from a given temporal data series of precipitation or temperature, the mean annual critical threshold exceedance frequency (i.e., TEF_a) was calculated as:

$$TEF_a = \frac{n}{Ndays} * 365 \qquad \text{Eq. 4.6}$$

where n is the number of events above the defined threshold and $Ndays$ represents the number of days with recorded data in the meteorological time series considered. The meteorological time series is either represented by the station data (with variable temporal coverage, see Section 4.3.3.1) otherwise by the time series recorded at each pixel of the grid-based meteorological dataset (spanning from 2003 to 2020). Independently from the input dataset, based on the climate normal concept, it is assumed that the available data, and therefore TEF_a, could be considered representative of the 30-year period 1991-2020 (World Meteorological Organization - WMO, 2007). Considering the different physical meaning and calculation procedure of each climatic index, the number of events above the critical threshold n was

derived accordingly. The calculation was carried out in the Matlab® environment, with the development of a specific code for each index.

The TEF_a for the EWI index was calculated on the P_{eff} daily dataset for the meteorological station time series. For the grid dataset the hourly precipitation time series aggregated at the daily time-step was used. Starting from a daily precipitation time series, the 1-, 3-, 7-, 15-, 30- and 60-day antecedent cumulated precipitation was calculated, in the same way as in Section 4.5.1.2. Secondly, it was verified if, for each day of the time series, the cumulated precipitation recorded in that day or in the antecedent days (i.e., 3 to 60 days) exceeded the threshold for at least a considered duration. Finally, each day was attributed a value of 1 or 0 for threshold exceedance or non-exceedance, respectively. At this point the parameter n in Eq. 4.6 for the EWI index was obtained in two ways:

(i) Given the series of 1 and 0, consecutive days recording threshold exceedance were counted as a single threshold exceedance event. This ensured their independency, with at least 24 hours of non-exceedance amongst them.
(ii) Given the series of 1 and 0, all the recorded 1 were counted as threshold exceedance events; retaining this information may be an indicator of the temporal "persistence" of above-the-threshold conditions.

From these two ways of interpreting the n parameter, two different TEF_a were obtained from Eq. 4.6, named EWIind (i.e., EWI independent) and EWIper (i.e., EWI persistence) respectively.

The TEF_a for the WD index was calculated on the original precipitation dataset with a 30-min time step for the meteorological station time series. For the grid dataset the hourly precipitation time series was used. Starting from a 30-min or hourly precipitation time series, for each time step in the time series (i.e., each 30 min or 1 hour), the number of WD episodes recorded in the antecedent 30-, 60-, 90-, 120-,180- and 365-day durations were calculated in the same way as in Section 4.5.1.2. Secondly, it was verified if, for the considered time series timestep, the number of episodes recorded in the antecedent days (i.e., 30 to 365 days) exceeded the threshold for at least one considered duration. Finally, each time step was attributed a value of 1 or 0 for threshold exceedance or non-exceedance, respectively. At this point, the parameter n in Eq. 4.6 for the WD index was simply obtained summing all the threshold exceedance events (i.e., the values equal to one in the data series). No further processing was needed to guarantee the independency of threshold exceedance events, as it was inherited by the definition of WD episodes in Section 4.5.1.2: two independent wet episodes must be separated by a dry period of at least 24 consecutive hours. Nevertheless, a correction to the number of events above the threshold for not-

heated rain gauges (i.e., $n_{not\text{-}heated}$) was adopted, since the time series used for the calculation did not considered snow melting inputs. The average percentage of threshold exceedance events occurred in the Winter period (i.e.,$\%\overline{n_{winter}}$) was calculated using the available heated rain gauges time series. Different Winter periods were considered for stations above and below 1500 m a.s.l.: November-April and December-March, respectively. This difference is due to the recorded inactivity periods (missing values) of stations at different elevations. Considering, for example, the available heated rain-gauges above 1500 m a.s.l., $\%n_{winter}$ was calculated with the following equation:

$$\%\overline{n_{winter}} = \frac{\sum_{i=0}^{S}\left(\frac{n_{nov-apr}}{n}100\right)_i}{S} \quad \text{Eq. 4.7}$$

where S is the total number of heated rain gauges above 1500 m a.s.l., n is the total number of events above the WD threshold and $n_{nov\text{-}apr}$ is the number of events above the WD threshold for the Winter months.

Subsequently, the final corrected n value for not-heated rain gauges (i.e., $n_{corrected}$) was obtained as:

$$n_{corrected} = n_{not-heated} + (\%\overline{n_{winter}} * n_{not-heated}) \quad \text{Eq. 4.8}$$

and used in Eq. 4.6 instead of n, when the TEF_a calculation involved a not-heated rain gauge.

The TEF_a for the FT index was calculated on the temperature dataset with a 30-min time step for the meteorological station time series. For the grid dataset the hourly temperature time series was used. For each time step in the time series (i.e., each 30 min or 1 hour), the number of FT episodes recorded in durations of 1-, 3-, 7-, 15-, 30-, 60-, 90-, 120-,180- and 365-day period were calculated in the same way as in Section 4.5.1.2. To ensure the independency of the threshold exceedance events, a more elaborated procedure than those for EWI and WD indices was applied, as FT cycles were defined only as transitions across 0°C. Each duration was treated separately and then aggregated. For the 1-day duration the procedure was rather simple (Figure 4.21).

The number of cycles accumulated in the 24 hours preceding the considered time-step were calculated and each time-step was attributed a value of “1” or “0” if the number of cycles exceeded or did not exceed the threshold, respectively. Subsequently, a threshold exceedance event was counted only if the 1 value were temporally distant at least 24 hours (i.e., 1 day).

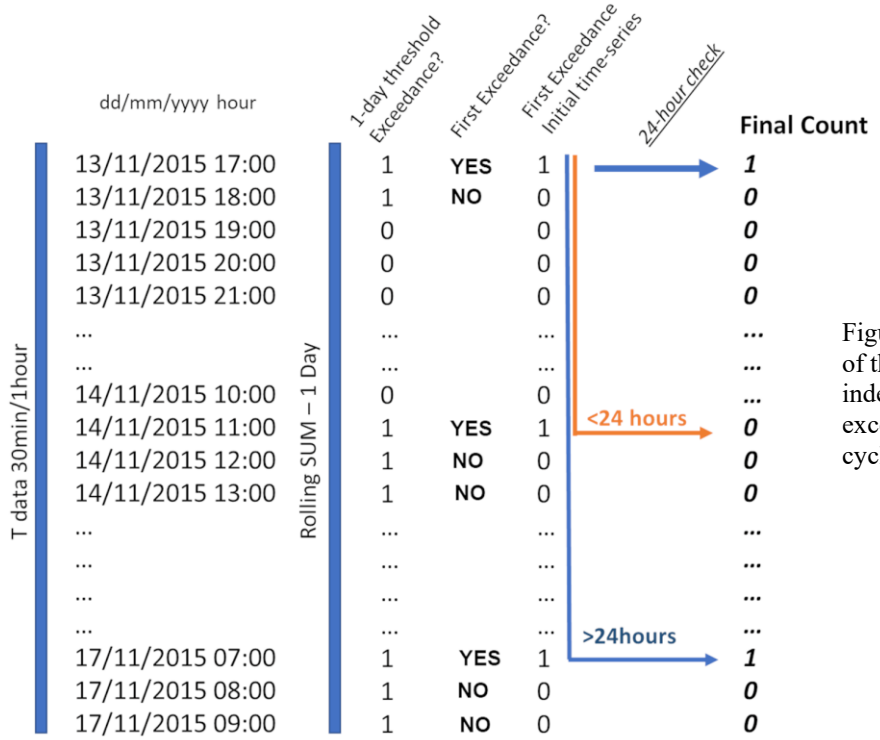


Figure 4.21 Schematic representation of the Matlab® code used to calculate independent 1-day threshold exceedance events for freeze-thaw cycles.

For 3-day to 365-day durations the procedure was slightly more articulated. A "cascade" incremental procedure was implemented, to ensure independency of threshold exceedance events amongst different durations. In particular, the secondary threshold exceedance events (and not only the first 1-value of a consecutive series) were taken into account too and expressed as "2". This was necessary to compare exceedances from different durations and potentially adding new independent exceedance events, otherwise not recorded. The 3-day procedure is explained hereafter and schematically presented in Figure 4.22.

The number of cycles accumulated in the 3-day period preceding the considered time step were calculated and each time step was attributed a value of "1" or "0" for threshold exceedance or non-exceedance, respectively. The 1-value were maintained if they represented the first recorded exceedance of a consecutive series of recorded "1", otherwise they were modified in a value of "2". The 2-value signifies that a threshold exceedance event was recorded, but it was already recorded in the antecedent time step. A time series of "0","1" and "2" was therefore obtained (i.e., the *exceedances initial time series*). From this series, a threshold exceedance event was counted only if: (i) the 1-value were temporally distant at least 72 hours (i.e., 3 days), and (ii) no threshold exceedance was already recorded, for the same specific time step, during the 1-day duration procedure. The result was an *independent first exceedances time series*, which was further compared to the *other exceedances initial time series*, with the aim to restore the 2-value. From such comparison, some "2" were counted as independent threshold

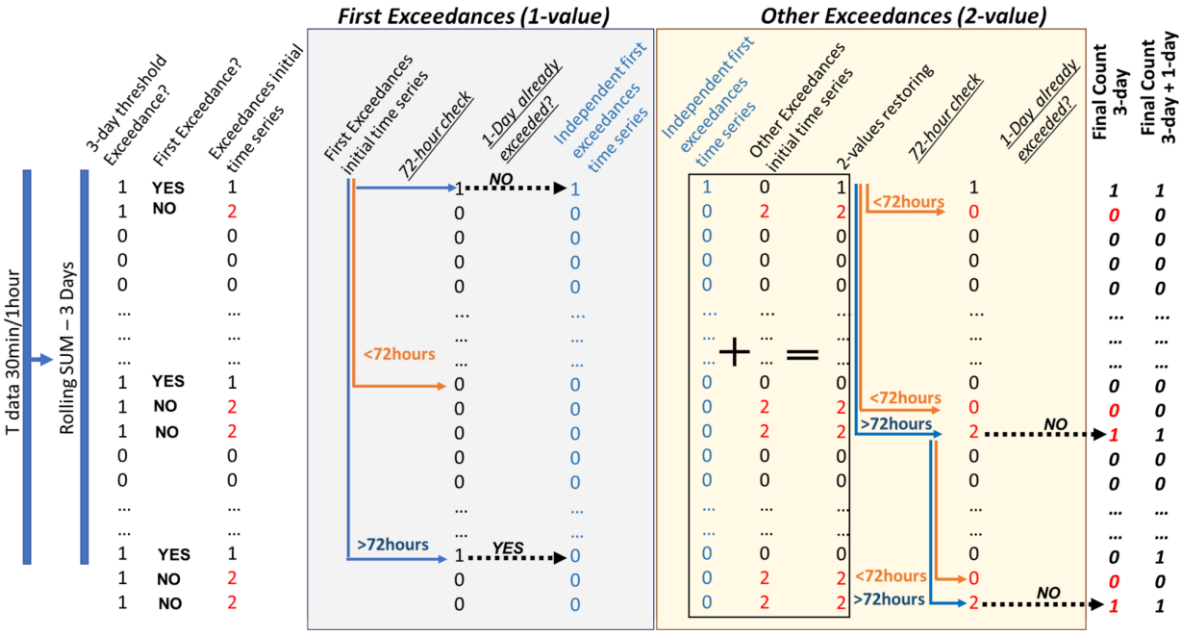


Figure 4.22 Schematic representation of the Matlab® code used to calculate independent 3-day threshold exceedance events for freeze-thaw cycles. Highlighted in red the not first exceedances.

exceedance events for the 3-days duration if: (i) they were temporally distant at least 72 hours from the last counted threshold exceedance events (both from 1-value in the *independent first exceedances time series* and from the previous counted 2-value) and (ii) there was no threshold exceedance recorded in that specific time step during the 1-day duration procedure.

The procedure iteratively continued until the longest duration (i.e., 365 days) with the same rationale and only with slight variations. Referring to Figure 4.22 the mechanism was the same, but the final count was the result of the comparison with all the previous durations shorter than the duration examined (and not only the 1-day duration as for the 3-day duration). Summarizing:

(i) For each duration considered, the temporal distance needed to ensure events independency would increase accordingly (e.g., for the 30-day duration, at least 720 hours of temporal distance).

(ii) Threshold exceedances for each duration were compared with *all* the previous durations (e.g., for the 30-day duration, independent threshold exceedances were maintained if no exceedances were already recorded in that specific time step for the previous 1-,3-,7- and 15-day duration).

(iii) Finally, the parameter n in Eq. 4.6 for the FT index was obtained summing all the independently counted threshold exceedance events.

For the EWI index, the threshold selected for the calculation of the mean annual exceedance frequency was the one normalized to RDN (refer to Section 4.5.2.3), as it is usually recognized as a better proxy

than the MAP for the occurrence of severe or extreme rainfall events (Wilson and Jayko 1997; Guzzetti et al., 2007). Furthermore, this normalization parameter is deemed more suitable to capture future climatic variations of the rainfall regime in the Alpine Region, expected to be characterized by an increasing frequency of extreme precipitation events (Rajczak et al., 2013; Gobiet et al., 2014; Ban et al., 2020). For the WD index, the threshold selected for the calculation of the mean annual exceedance frequency was the global threshold defined in Section 4.5.2.4. For the FT index, the threshold normalization to the FTN parameter did not produce any substantial improvement in comparison to the non-normalized one (see Section 4.5.2.5). Thus, it was not possible to establish a best threshold between the two. Accordingly, both for meteorological stations and for the grid dataset the mean annual exceedance frequency for FT was calculated using both.

The TEF_a are valid for the reference period 1991-2020 and are potentially modifiable and updatable for future reference periods, exploiting climate projections of precipitation and temperature. Difference maps for each index were also created in order to compare the spatial distribution of the TEF_a coming from the station-based and grid-based datasets.

4.6.1.2 Station-based climatic predictors regionalization

Before their use in the susceptibility analysis, the station-based TEF_a of each index were regionalized. At this scope, different deterministic (IDW, TPS), geostatistical (OK, KED) and regression (MLR, GWR) techniques were evaluated. Details on the regionalization techniques can be found in Section 2.2.

Each technique was applied testing several model parameter and environmental covariates (e.g., elevation, latitude, longitude, slope and aspect) combinations. For IDW, different powers were tested (i.e. 1.5, 2, 2.5, 3, 3.5, 4,4.5, 5). For geostatistical techniques, different variogram models were applied (e.g. exponential, circular, spheric, gaussian). For TPS either x-y coordinates alone or in combination with altitude, slope or aspect were tested as covariates. For GWR all the possible model parameters (i.e., fixed or adaptive bandwidth, gaussian or bisquare kernel and CV or AIC method) and covariates combinations (one to all covariates) were tested. The analyses were performed in the R environment (https://www.r-project.org/) by means of the libraries *gstat* (Pebesma, 2004), *fields* (Nychka et al., 2017) and *spgwr* (Bivand and Yu, 2021).

Model performances were assessed through a leave-one-out cross-validation (LOO-CV) calculating goodness-of-fit and error metrics (correlation coefficient r, and NMRSE). For geostatistical techniques, starting from the selected variogram model on the entire dataset (i.e., “general model”), range and sill were manually fitted on each cross validation set. Since the station number was limited, the manual selection of the variogram models during the LOOV-CV had allowed a further control during the

regionalization procedure. To verify the consistency and the maximum variation from the general model, histogram of sill and range coming from the LOO-CV were produced. Finally, for each property, best results coming from each technique are presented and compared, and the maps considered as the most reliable were selected as the climate-related predictors for the susceptibility model. To limit collinearities between predictors, in case the regionalization performance of two or more techniques was the same for a specific index, the one with less covariates was preferred. In addition, a check of the predicted minimum and maximum values, in comparison to the real TEF_a values at the stations, was carried out for all regionalization outputs to exclude maps characterized by unreliable values (i.e., negative values) and excessive extrapolation, i.e., more than 90% higher (lower) than the maximum (minimum) values recorded at the station.

4.6.1.3 Grid-based climatic predictors aggregation

The TEF_a of each index were calculated on a pixel basis, introducing in the Matlab® codes developed in Section 4.6.1.1 functions and processes for parallel computing. In particular, the *parfor* function of the Parallel Computing Toolbox® was employed in order to speed up the calculations. The datasets (with dimensions>200 GB each) were treated as three-dimensional matrices, in which each pixel may be viewed as a weather station (Figure 4.23). The total calculation time in parallel mode was 53 hours for FT, 43 hours for WD and 40 hours each for EWIind and EWIper.

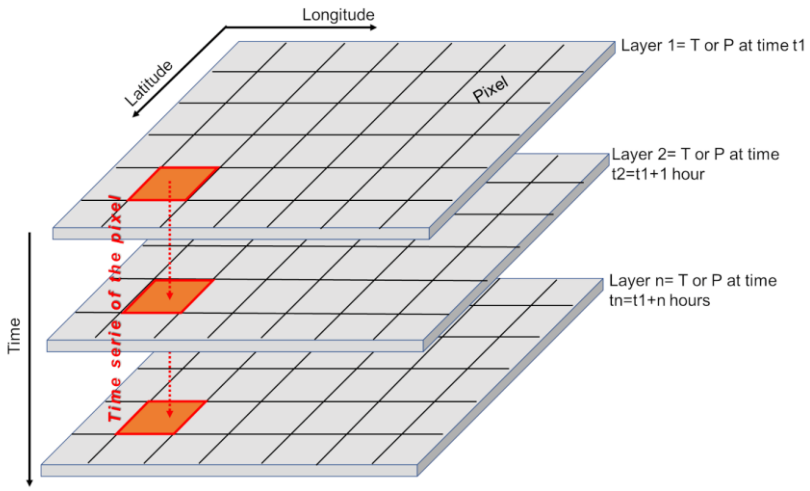


Figure 4.23 Schematic representation of the pixel-by-pixel calculation of grid-based climatic predictors.

4.6.1.4 Snow melting predictors creation

Two summary predictors related to snow melting dynamics were introduced in the susceptibility model, namely SWEep (the average number of melting events occurring over 16-day periods in a hydrological year) and SWEmax (the maximum amount of melting recorded over 32-day periods in the whole data series). These predictors were developed by Camera et al. (2021) in the same study area for a shallow

landslide susceptibility model, starting from the SWE gridded dataset available for the study area (see Section 4.3.3.3). Firstly, the SWE pixel values were aggregated in sub-basins, to overcome the issue related to the SWE dataset's lack of mass conservation to the single cell. Secondly, the consistency between the snow dynamics reproduced by the raster and the meteorological station datasets was verified by means of correlations calculating R^2 as an index of the fit (additional details in Raffa, 2020 and Camera et al., 2021).

4.6.2. Results

The TEFa calculated for each meteorological station are presented in Table 4.4, while the best regionalized maps are presented in Figure 4.24. For EWIind (Figure 4.24a), the best regionalization results were obtained from the application of a MLR technique with altitude and longitude as covariates (r= 0.72 and NMRSE=0.186). For EWIper (Figure 4.24b), IDW, MLR with the covariates altitude and slope and KED with the covariates altitude, slope and longitude, resulted in the same quantitative performance (i.e., r=0.66 and NMRSE=0.21). Thus, the IDW derived map was selected as it did not include covariates.

Table 4.5 reports the regionalization results deriving from all the tested techniques for EWIind and EWIper, which showed an opposite behaviour in terms of altitude dependence, with an indirect and direct relationship respectively. At high altitudes, TEF_a for EWIper showed the highest values, indicating temporally clustered threshold exceedance events (i.e., a higher number of consecutive days in above-the-threshold conditions than at low altitudes). Conversely, TEF_a for EWIind showed its highest values at low altitudes, indicating that independent threshold exceedance events are more frequent but shorter than at high elevations.

For WD (Figure 4.24c,d), the regionalization deriving from the application of KED with the latitude covariate was selected (r=0.65 and NMRSE=0.23). The variogram was best fit by an isotropic spherical model with a nugget effect of 50, a partial sill of 740 and a maximum range of 7 km. Even if TPS, MLR and GWR techniques performed slightly better than KED, they led to very low minimum predicted values of threshold exceedance frequency (TEF_a =1, while the minimum value recorded at the stations was 33), considered not reliable. Table 4.6 reports the regionalization results deriving from all the tested techniques for WD. The spatial pattern of the WD map (Figure 4.24c) revealed that the northern part of the study area, and in particular the head of the Valtournenche Valley, is more strongly affected by threshold exceedance events in terms of wet and dry episodes than the E-W valley axis and the southern part of the study area.

Table 4.4 Mean Annual Threshold Exceedance Frequency (TEF_a) for EWIind, EWIper, WD and FT calculated at each meteorological station . For FT, TEF_a were calculated both starting from the not-normalized threshold (TEF_a not-norm) and from the threshold normalized to the FTN parameter (TEF_a norm).

Station ID	*Ndays*	***EWIind***		***EWIper***		***WD***				***FT***			
		n	TEF_a	*n*	TEF_a	*Ndays*	*n*	$n_{corrected}$	TEF_a	*Ndays*	*n*	TEF_a *(not-norm)*	TEF_a *(norm)*
ST13	4243	219	19	713	61	3982	841	1001	92	9130	515	21	38
ST17	10592	513	18	1397	48	6209	787	787	46	4391	176	15	30
ST28	6209	287	17	878	52	6209	567	567	33	6208	299	18	34
ST4	4193	200	17	710	62	4207	628	628	54	4191	261	23	26
ST16	2988	148	18	423	52	2988	315	375	46	2986	172	21	30
ST3	6209	265	16	1005	59	6209	676	676	40	6208	196	12	21
ST9	6879	288	15	1354	72	6209	1056	1257	74	6877	1046	56	68
ST10	4185	158	14	983	86	5627	391	528	34	5625	623	40	63
ST14	5966	235	14	963	59	5966	1684	1684	103	5965	462	28	43
ST26	9496	357	14	1803	69	-	-	-	-	-	-	-	-
ST5	3128	115	13	768	90	3128	1167	1167	136	3126	514	60	68
ST1	6110	217	13	976	58	6209	1359	1359	80	6108	1011	60	79
ST27	10408	412	14	1991	70	-	-	-	-	-	-	-	-
ST11	4177	134	12	800	70	4177	342	462	40	4175	686	60	75
ST12	4186	123	11	811	71	4186	587	792	69	4184	702	61	75
ST8	10592	382	13	1957	67	5627	1912	1912	124	5626	1156	75	89
ST7	4401	92	8	1129	94	4328	725	979	83	4400	1025	85	85
ST2	-	-	-	-	-	4108	517	615	55	-	-	-	-
ST1bis	-	-	-	-	-	6209	1359	1359	80	-	-	-	-
ST14bis	-	-	-	-	-	5966	1684	1684	103	-	-	-	-
ST6	-	-	-	-	-	-	-	-	-	5602	1212	79	71

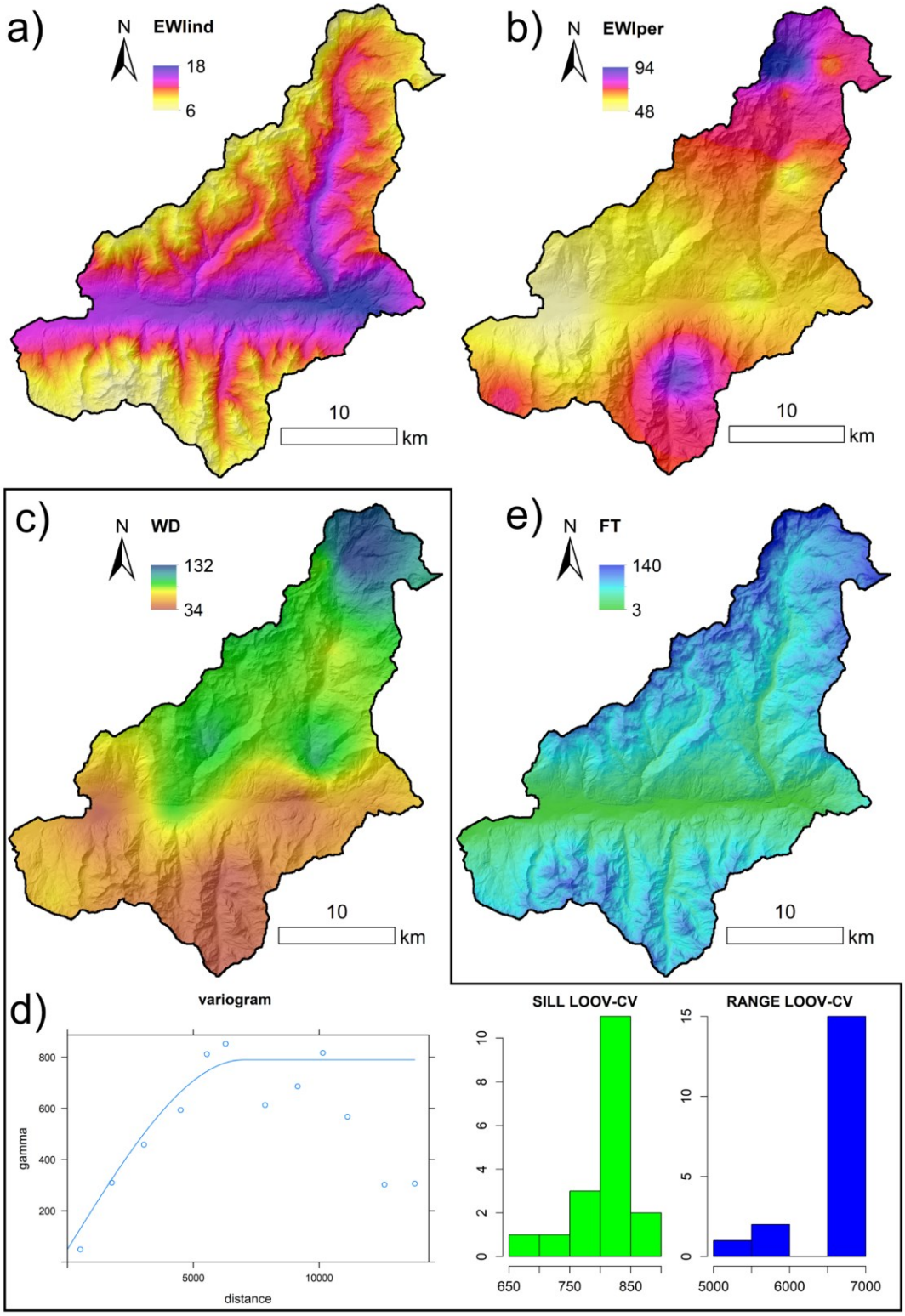


Figure 4.24 Final maps for station-based a) EWIind; b) EWIper; c) WD, with the d) associated variogram and sill and range variations from the LOOV-CV runs; e) FT.

Table 4.5 Performance regionalization results (r correlation coefficient and Normalized Mean Root Square Error-NMRSE) for EWIind and EWIper.

Technique	Parametrization	r correlation coefficient observed-simulated	NMRSE
	EWIind		
IDW	Idp=2	r<0	
TPS	Covariate=coordinates+elevation	0.65	0.21
MLR	Covariate=elevation+longitude	0.72	0.18
GWR	Bisquare kernel, fixed bandwitdth, CV criterion Covariate=elevation+longitude	0.66	0.20
OK	No variogram structure found		
KED	Spheric Isotropic Lag=1600 m Covariate=elevation+longitude	0.68	0.19
	EWIper		
IDW	Idp=2	0.66	0.21
TPS	Covariate=coordinates+elevation	0.36	0.29
MLR	Covariate=elevation+slope	0.66	0.21
GWR	No bandwidth found		
OK	No variogram structure found		
KED	Spheric Isotropic Lag=1200 m Covariate=elevation+longitude+slope	0.66	0.21

Table 4.6 Performance regionalization results (r correlation coefficient and Normalized Mean Root Square Error-NMRSE) for WD.

Technique	Parametrization	r correlation coefficient observed-simulated	NMRSE
	WD		
IDW	Idp=2	0.55	0.25
TPS	Covariate=coordinates+aspect	0.69	0.22
MLR	Covariate=aspect+latitude	0.70	0.22
GWR	Bisquare kernel, fixed bandwitdth, CV criterion Covariate=aspect+latitude	0.67	0.22
OK	No variogram structure found		
KED	Spheric Isotropic Lag=1200 m Covariate=latitude	0.65	0.23

For FT (Figure 4.24e), it was verified that the correlation (in terms of r and p-value) between the TEF_a obtained from the two thresholds (the not-normalized and the normalized one) was strong and

significant. With a correlation coefficient of r=0.95 and a p-value<0.005, the two thresholds can be considered interchangeable and expressing the same process. Thus, for the station derived predictor, it was decided to adopt the TEF_a deriving from the simpler not-normalized threshold and proceed with its regionalization step. The best regionalization results were obtained from the application of an MLR technique (r=0.93 and NMRSE=0.12) with the covariates altitude, latitude and slope. The GWR technique with the same covariates returned an equivalent performance. However, as for some CV runs the bandwidth converged to the upper bound (i.e., the GWR local domain of validity coincides to the MLR domain validity), the MLR technique was adopted. Table 4.7 reports the regionalization results deriving from all the tested techniques for FT.

Table 4.7 Performance regionalization results (r correlation coefficient and Normalized Mean Root Square Error-NMRSE) for FT.

Technique	**Parametrization**	**r correlation coefficient observed-simulated**	**NMRSE**
	FT		
IDW	Idp=2	0.78	0.2
TPS	Covariate=coordinates+elevation	0.90	0.146
MLR	Covariate=elevation+latitude+slope	0.93	0.12
GWR	Gaussian kernel, fixed bandwitdth, CV criterion Covariate=elevation+latitude+slope	0.93	0.12
OK	No variogram structure found		
KED	Exponential Isotropic Lag=2000 m Covariate=elevation+latitude	0.92	0.13

The spatial pattern of threshold exceedance frequency showed a strong, direct relationship with altitude. However, expected temperature inversion effects or the individuation of the highest threshold exceedance values at mid-high-altitude zones below the freezing level, observed between 3850 and 4200 m a.s.l. for Aosta Valley (Ponziani et al., 2020), were not revealed by the FT regionalization. This was mainly imputable to the fact that the regionalization relies on the strongest correlation signal, which in this case was the direct linear relationship among the FT threshold exceedance frequency and altitude. The same issue could be expected deriving the TEF_a through the regionalization of the normalized threshold.

The grid-based TEF_a maps of each index, calculated on a pixel basis, are presented in Figure 4.25. EWIind spatial distribution (Figure 4.25a) did not show the same strong direct correlation with altitude as observed in the station-derived map. Rather, it showed the highest values concentrated along the entire

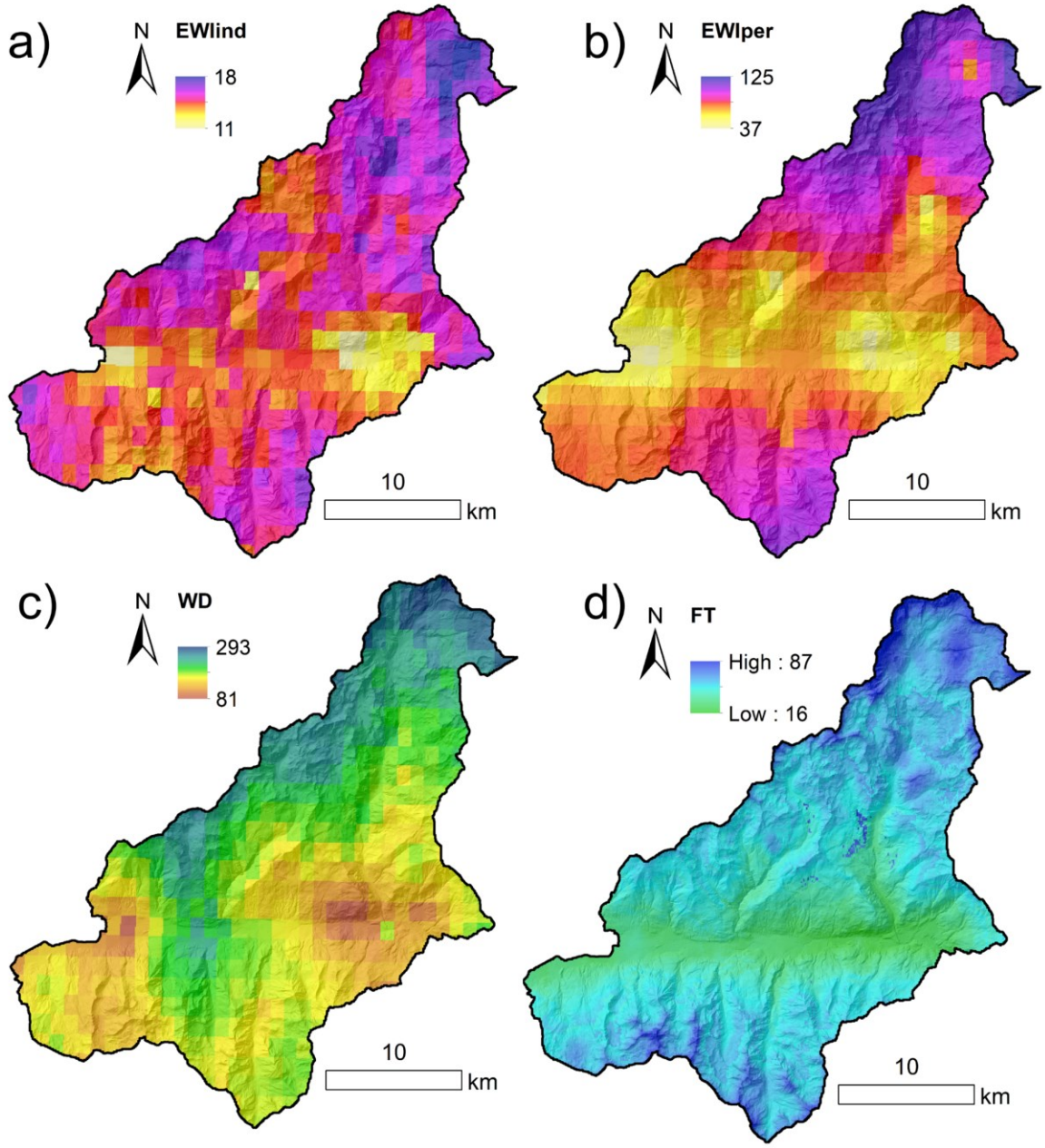


Figure 4.25 Final grid-based maps for a) EWIind, b) EWIper, c) WD and d) FT.

Valtournenche N-S axis. Conversely, the highest EWIper values (Figure 4.25b) were located at the highest elevation, as observed in the station-derived maps. The calculated value ranges were quite similar to those of the regionalized maps for EWIind, except for the minimum values which resulted higher (approximately 45%) for the grid-based maps. Regarding EWIper, maximum values were higher (approximately 25%) for the grid-based maps than for the station-based ones, whereas minimum values were lower (approximately 23%) for the grid-based map.

The maps showing the above discussed differences in the spatial distribution between the grid-based and the station-based derived maps for EWIind and EWIper are presented in Figure 4.26a and Figure 4.26b.

The grid-derived map of WD (Figure 4.25c) showed a spatial pattern similar to the station-based WD map; the head of the Valtournenche Valley had the highest threshold exceedance frequency while the E-W Valley axis the lowest one. However, the grid-based map also individuated a belt of high values in the middle of the southern slope of Mount Emilus Mountain Community and along the NW border of the study area, both not observable in the station-based map. Moreover, the minimum and maximum values of threshold exceedance for WD obtained by the grid-based maps were much higher than in the station-based maps. The map showing the differences in the spatial distribution between the grid-based and the station-based derived maps for WD is presented in Figure 4.26c.

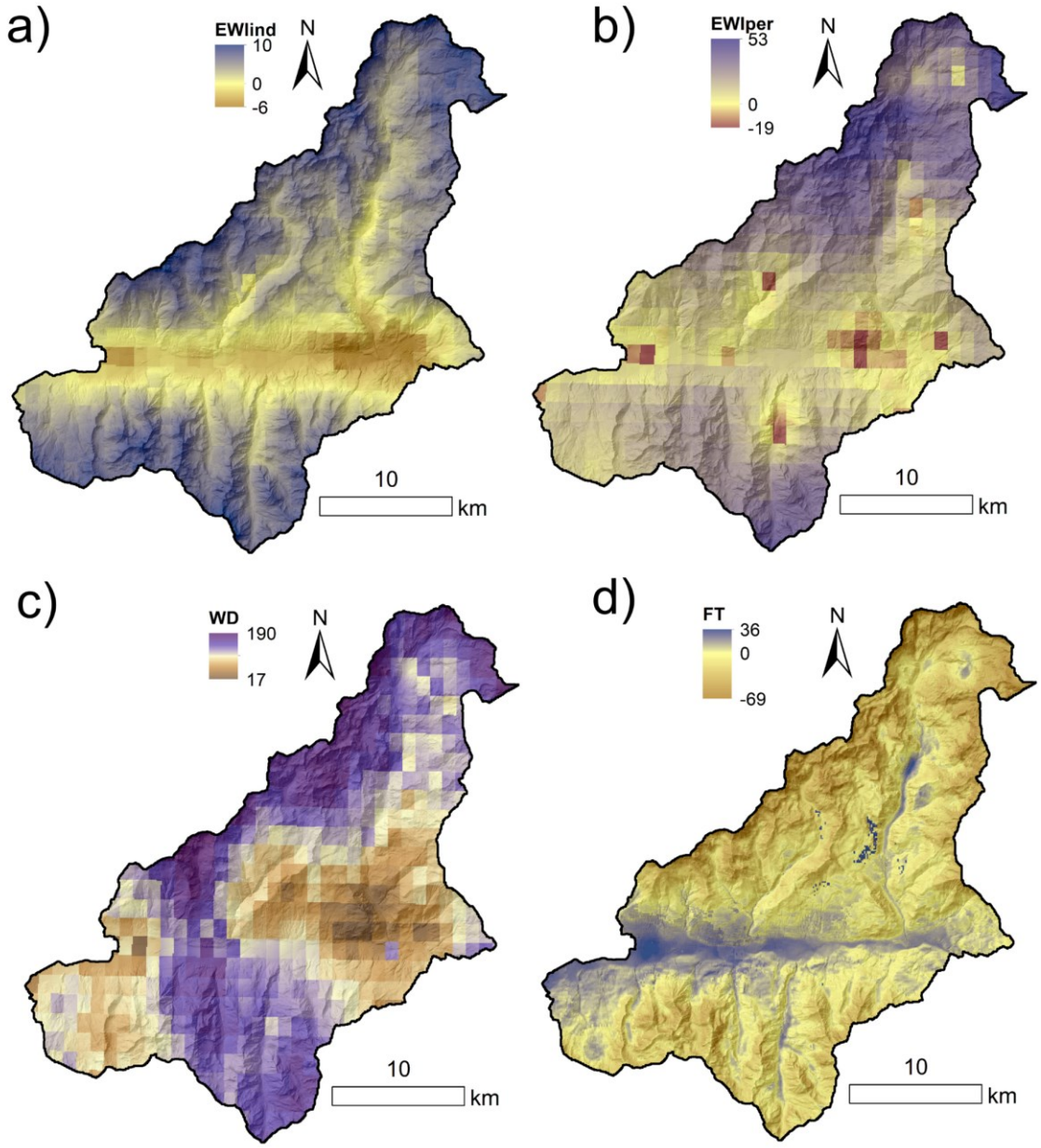


Figure 4.26 Maps showing the spatial pattern of the differences between the grid based and station based climatic predictors for a) EWIind, b)EWIper, c) WD and d)FT.

The pixel-based comparison between the TEF_a obtained from the normalized and non-normalized FT thresholds resulted in a correlation coefficient of r=0.87 and a p-value<0.005. This is consistent with the results obtained for the station-based FT maps and demonstrates that the two thresholds are interchangeable. Nevertheless, to test them both in the susceptibility model it was arbitrarily decided to adopt the TEF_a deriving from the normalized threshold for the grid derived FT predictor. For the grid-derived map of FT (Figure 4.25d), the spatial pattern strongly correlated with altitude was preserved. However, it was possible to observe some bands and zones at mid-high altitudes where the threshold exceedance frequency has very high values, close to the maxima. The grid dataset was therefore suitable to capture patterns attributable to temperature inversion phenomena and the seasonal shifting of the zero isotherm; these areas are the positive values represented by blue shades in Figure 4.26d. The positive values in Figure 4.26d (i.e., the brown shades) indicated areas where the threshold exceedance frequency was higher for the station-derived FT. The usage of two slightly different thresholds might have had a role in causing these differences, too.

Regarding snow melting predictors, the SWEmax and SWEep maps from Camera et al. (2021) are presented in Figure 4.27. Note that for SWEmax, since snowmelt is a depletion of the SWE storage, the values are negative.

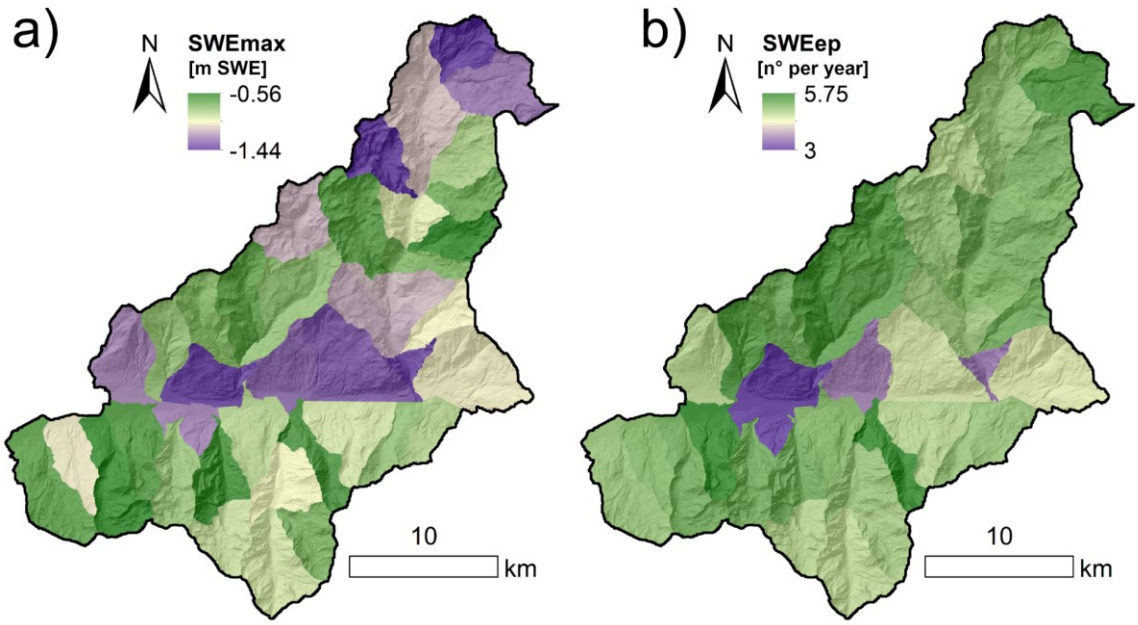


Figure 4.27 Snow melting predictors included in the susceptibility modelling phase, from Camera et al., 2021. a) SWEmax, i.e., the maximum amount of melting recorded over 32-day periods in the whole data series and b) SWEep i.e., the average number of melting events occurring over 16-day periods in a hydrological year.

4.7. Rockfall susceptibility modelling

4.7.1. Methods

4.7.1.1 Traditional geo-environmental predictors

Besides the synthetic climatic predictors presented in Section 4.6, the traditional DEM-derived predictors (Van Westen, 2008; Reichenbach et al., 2018) were included in the rockfall susceptibility model. They were elevation, slope, aspect - included as northness = cos(aspect) and eastness = sin(aspect) -, profile curvature, plan curvature, and SAGA Topographic Wetness Index (SWI). They were derived at a 10 m x 10 m horizontal resolution, using the *RSAGA* package (Brenning et al., 2018). Geology was introduced as a categorical predictor with four classes, based on the reclassification of the geological map (1:10000 scale) available for the study area (Section 4.3.2). The distinction adopted followed a combined lithological and metamorphic criterion, as besides the strength characteristics imputable to the different lithologies, the widely variable metamorphic imprint characterizing the study area (see Section 4.2) may further differentiate the rock mass properties. The classes were: (i) Zermatt-Saas ophiolites in the eclogitic facies, (ii) Austroalpine upper outliers in the blueshist facies, (iii) Combin ophiolites in the blueshist facies, (iv) Austroalpine lower outliers in the eclogitic facies. A more detailed distinction based on peculiar lithologies and mineralogical associations as the one adopted in Section 4.5.1.6 (with 10 categories) was not suitable for the susceptibility modelling phase as it comprises too numerous categories to be handled by the model, possibly causing problems during the model validation step (Guzzetti et al., 2006). The reference class adopted as the modelling reference class was the most abundant in the study area (the Zermatt Saas ophiolites).

4.7.1.2 Rockfall susceptibility model setup and assessment

The 243 rockfall records occurred between 1990 and 2018 (both precisely dated and not precisely dated, see Section 4.3.1) were used as the rockfall "presence" inputs of the binary response variable for the susceptibility modelling phase. The availability of the date of occurrence (or period of occurrence) for the rockfall inventory was a crucial characteristic for the inclusion of climate-related predictors in the susceptibility model. Rockfall absence locations were selected as for the Valchiavenna Case Study (Chapter 3), by a random extraction from the "eligible" area. The "eligible area" was obtained from a preliminary masking of urban areas, glaciers, water bodies and quaternary deposits (i.e., not-modellable areas) as mapped in the geological-geomorphological map at the 1:10000 scale presented in Section 4.3.2. Also, areas within a 100-m buffer from rockfall points (i.e., non-eligible areas) were excluded. This buffer dimension was selected after an analysis of the mean (138 m) and median (108 m) radius of 59 scarps available in the above-mentioned geological-geomorphological map, considered

representative of the geometrical characteristics of rockfall processes in the area. Besides, it resulted very similar, and thus consistent, to the one used for the Valchiavenna Case Study (80 m). Although the geomorphological dataset was not suitable for rockfall susceptibility modelling (lack of date of occurrence, not univocal ID associating scarps and rockfall deposits) it was still useful as it added information in comparison to the dated inventory. In the latter, being point based, details on the size of the rockfall detachment areas were missing. A 1:1 ratio was adopted for the extraction of absence points, as the percentage of the "eligible area" on the "modellable area" was 97.6%, following the guidelines of Hong et al. (2019) who found that this ratio is optimal when the "eligible" area approaches the 99% of the "modellable area". The locations of rockfall and non-rockfalls (i.e., negative) points and the "eligible" area are shown in Figure 4.28.

Rockfall susceptibility modelling was performed using Generalized Additive Models (GAMs), with the procedural scheme adopted in Chapter 3 (refer to Section 3.4.5.1) through the R package *mgcv* (Wood, 2017) and summarized hereafter. Variable selection was performed through the *shrinkage* option,

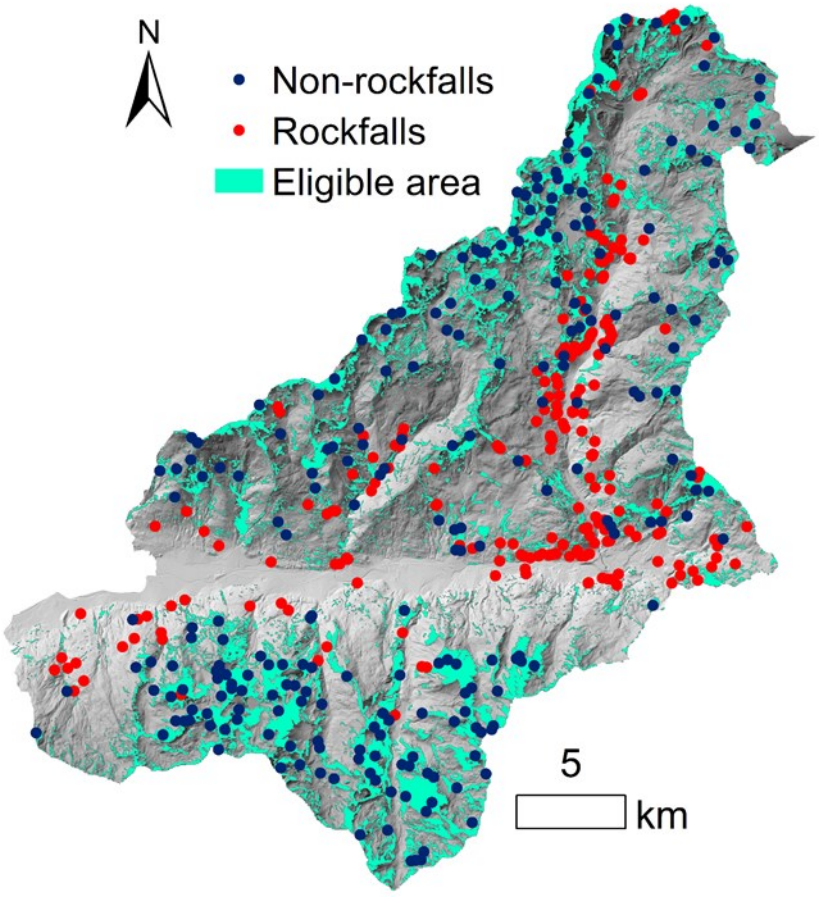


Figure 4.28 The locations of rockfall and non-rockfalls (i.e., negative) points and the "eligible" area.

available in the *mgcv* package, while predictors behaviour was analysed through the associated smoothing functions by means of CSF plots and odds ratios for geology categories. A repeated, k-fold cross-validation framework was configured to estimate the AUROC; differently from the Valchiavenna case study, model performance was assessed through both a non-spatial (nsCV) and a spatial (sCV) k-fold cross-validation, with a random spatial partitioning and a partitioning based on k-means clustering of coordinates, respectively. Both cross-validation types were set up with $k=5$ folds and r=100 repetitions, to obtain results independent from a specific partitioning. The cross-validation was implemented using the R package *sperrorest* (Brenning, 2012a). Predictors' smoothing functions from each cross-validation run were compared with the corresponding smoothing function obtained on the entire dataset, to assess coherency and robustness of their behaviour. To assess the importance of predictors, the penalization frequency coming from the application of the shrinkage option was considered. It was calculated as the percentage of CV runs in which the effective degrees of freedom, *edf*, resulted much lower than 1 (here a 0.7 threshold was selected). Penalization frequency was then combined with the mean decrease in deviance explained (mDD%), calculated as in Knevels et al. (2020). Moreover, concurvity between the smoothers, i.e., the generalization of multicollinearity to non-parametric functions, was calculated.

4.7.1.3 Four steps to increase the physical plausibility of the statistical model

Recently, some researchers have conveyed interest in depicting and discussing the effect of inventory collection procedures and deriving bias, on predictors behaviour physical plausibility and spatial inconsistencies in the resulting susceptibility map (Steger et al., 2016a, 2016b, 2017; Jacobs et al., 2020; Camera et al., 2021; Bajni et al., 2021b). Some works were focused on incorporating strategies in the modelling procedure aimed at limiting the inventory bias effects (Bornaetxea et al., 2018; Knevels et al., 2020; Steger et al., 2021), which may threaten results reliability and usability. Finally, it is a common practice to carry out a multicollinearity diagnosis of predictors before using them as independent variables in a susceptibility model, both when dealing with logistic regression and complex Machine Learning models (Chen W. et al., 2017; Nohani et al., 2019; Roy et al., 2019; Yi et al., 2020; Arabameri et al., 2020; Alqadhi et al., 2021). In the case of climatic predictors, multicollinearity and concurvity issues between each other and with morphometric predictors have to be expected and consequently handled. To do so, a four-step modelling procedure was adopted (Figure 4.29). For all steps, the modelling algorithm and the model evaluation procedure remained fixed as described in the previous section.

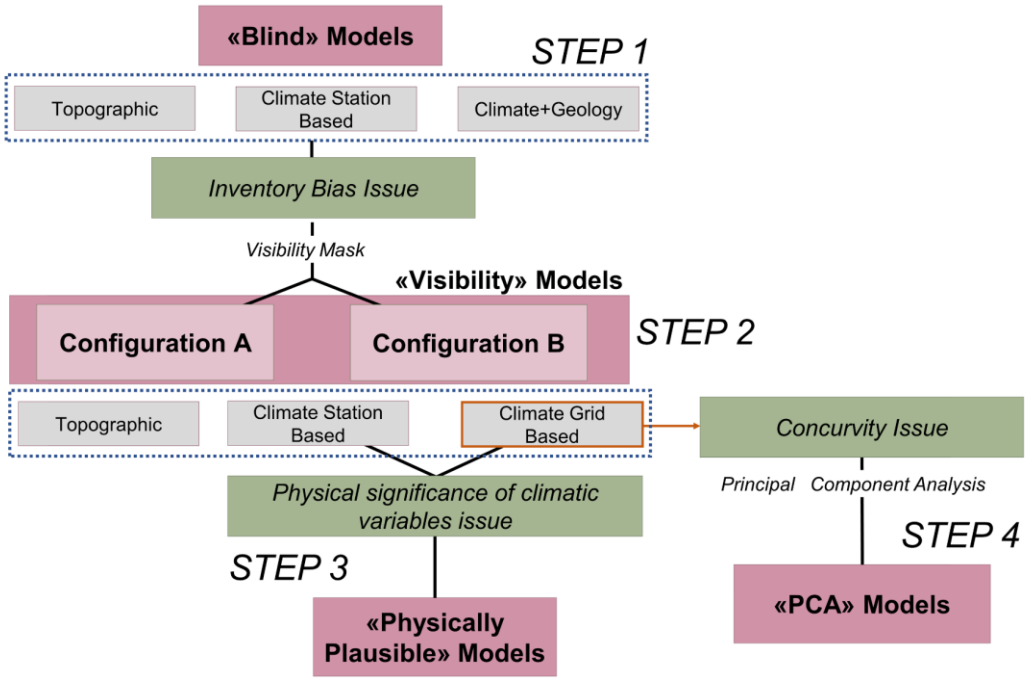


Figure 4.29 Workflow for the four-steps adopted susceptibility modelling procedure.

The first step was entirely dedicated to the "Blind Models", which are those that ignore, on purpose, the potential implication of an incomplete or biased rockfall inventory. Steger et al. (2021) observed that the majority of landslide susceptibility related works usually disregard inventory related issues, relying exclusively on quantitative performance assessment and neglecting the geomorphological and physical plausibility of the results. Three "Blind" models with different predictors sets were produced. Their characteristics are presented in Table 4.8.

Table 4.8 Model ID and predictors set used in the three "Blind" models produced. Note that the Blind-CLIMATE and Blind-GEO were constructed using the not-penalized morphometric predictors from the Blind-TOPO model.

Model ID	***Morphometric predictors***	***Climatic predictors***	***Geology***
Blind-TOPO	Elevation, slope, northness, eastness, SWI, profile curvature, plan curvature	-	-
Blind-CLIMATE	Not-penalized predictors from Blind-TOPO	Station-based *EWInd,WD,FT SWEep SWEmax*	-
Blind-GEO	Not-penalized predictors from Blind-TOPO	Station-based *EWInd,WD,FT SWEep SWEmax*	Reclassified geology

The second step was aimed at reducing the inventory bias, which causes an overrepresentation of rockfalls close to infrastructures and roads, thus considering the data collection procedure. Firstly, it was assumed that rockfall event reporting is dependent on the reachability of the sites and on the presence of infrastructures or other points of interest (Figure 4.30a). Principal and secondary roads shapefiles were obtained from the Open Street Map dataset (https://www.openstreetmap.org/#map=13/45.7400/7.4477), while the other points of interest were obtained from the publicly available PTP (Piano Territoriale Paesistico) of the Aosta Valley Region (https://geoportale.regione.vda.it/download/ptp/). Main roads are located in the principal valleys bottom and it was assumed that these roads are involved in a frequent and ordinary monitoring activity. Secondary roads comprehend connections between the principal valley towns and isolated villages at high altitudes, service roads used for dams, hydroelectric power plants and ski areas maintenance activities, forestry and agricultural roads, private roads etc. Among these typologies, only asphalted secondary roads or secondary roads leading to points of interest, such as ski areas, dams and hydroelectric power plants, meteorological stations and geo-cultural sites (e.g., the Saint Marcel ancient mining site), were retained (Figure 4.30c).

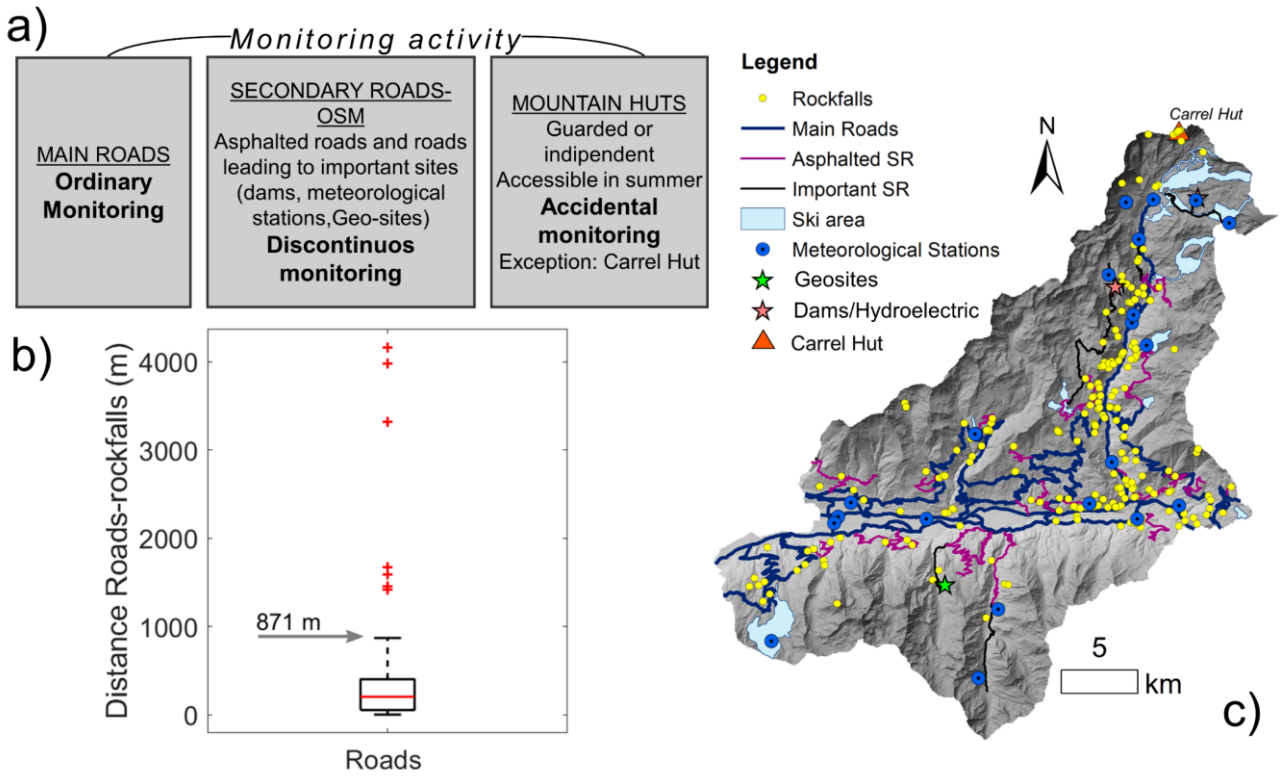


Figure 4.30 a) Monitoring activity rationale; b) boxplot representing the distribution of the rockfalls distances form principal and secondary roads; c) map reporting principal and secondary roads and principal infrastructures in the study area.

It was assumed that these roads represented less frequently beaten tracks and thus the monitoring activity is discontinuous but still effective. At this point, the distance between each rockfall and the closest roads (principal or secondary indifferently) was calculated, and the associated statistical distribution represented in the form of a boxplot (Figure 4.30b).

Subsequently, following a similar approach as in Boarnextea et al. (2018) and Knevels et al. (2020), the Visibility tool of ESRI ArcGIS® 10.2.2 was adopted to extract a visibility mask in which, reasonably, the monitoring designated personnel may record rockfalls as they are visible along the driving routes. The upper whisker of the boxplot (i.e., the largest value of the distribution that is no greater than the third quartile plus 1.5 times the interquartile range and corresponding to 871 meters) was selected as the maximum visibility parameter required by the algorithm. Another type of monitoring ("accidental" monitoring) may occur when high elevation rockfalls are observed by hut managers or hikers. It was deemed not appropriate to include this type of infrastructures in the scheme, except for the Carrel Hut, positioned on the southern slope of Matterhorn. This hut is characterized by the presence of a micro-seismic monitoring system (Amitrano et al., 2010; Occhiena et al., 2012) and thus assumed frequently visited, also for technical-scientific purposes. A buffer centred in the Carrel hut location was added to the visibility mask produced for roads, in order to include rockfall events in its surroundings. The buffer distance was chosen to maximize the number of includable rockfalls, and contemporarily to not overcome the distance selected for roads (i.e., 871 m). A total of 9 rockfalls were therefore retrieved within a maximum distance of 761 m from the hut.

The final visibility mask was also used to reduce the "eligible" area for non-landslide points selection. Therefore, a new non-rockfall points random selection was carried out following two slightly different approaches, which led to two different modelling configurations:

- Configuration A: out of the rockfall population of 243 events, 42 events (i.e., approximately the 17% of the inventory) remained outside the visibility mask and thus were excluded from the model training. The remaining 201 rockfalls were associated to an equal number of non-rockfalls points, randomly selected inside the visibility mask to produce a 1:1 ratio. The rockfalls excluded were used as a holdout independent test.
- Configuration B: to the 201 non-rockfall points already selected for Configuration A, 42 non-rockfall points were randomly selected outside the visibility mask, to match with a similar frequency the previously excluded 42 rockfalls and use the entire inventory.

In both cases, the model was trained in the eligible areas, but the predictions are extended to the whole study area. For each configuration, five "Visibility" models were produced, and their characteristics are

summarized in Table 4.9. It is important to outline that it was not possible to include geology in the modelling procedure: by further limiting the study area with the aim of bias reduction, two categories out of four remained without or with few sampling points, leading to the non-convergence of the model during the cross-validation process. This is a known issue, especially for categorical variables as rock type and land cover, which hampers the possibility to transfer the trained model from the training to a test area (Guzzetti et al., 2006).

Table 4.9 Model ID and predictors set used in the ten "Visibility" models produced, five for configuration A and five for configuration B. Note that all the models containing climatic predictors were constructed using the not-penalized morphometric predictors from the TOPO-A and TOPO-B models. Each of the two configurations were constructed either with station based climatic predictors (alternatively using EWIind or EWIper) or with grid based climatic predictors (alternatively using EWIind or EWIper).

Configuration	***Model ID***	***Morphometric predictors***	***Climatic predictors***
A	TOPO-A	Elevation, slope, northness, eastness, SWI, profile curvature, plan curvature	-
	ST Mode1A	Not-penalized predictors from TOPO-A	Station based *EWIind, WD, FT SWEep, SWEmax*
	ST Model2A	Not-penalized predictors from TOPO-A	Station based *EWIper, WD, FT SWEep, SWEmax*
	GR Model1A	Not-penalized predictors from TOPO-A	Grid based *EWIind, WD, FT SWEep, SWEmax*
	GR Model2A	Not-penalized predictors from TOPO-A	Grid based *EWIper, WD, FT SWEep, SWEmax*
B	TOPO-B	Elevation, slope, northness, eastness, SWI, profile curvature, plan curvature	-
	ST Mode1B	Not-penalized predictors from TOPO-B	Station based *EWIind, WD, FT SWEep, SWEmax*
	ST Model2B	Not-penalized predictors from TOPO-B	Station based *EWIper, WD, FT SWEep, SWEmax*
	GR Model1B	Not-penalized predictors from TOPO-B	Grid based *EWIind, WD, FT SWEep, SWEmax*
	GR Model2B	Not-penalized predictors from TOPO-B	Grid based *EWIper, WD, FT SWEep, SWEmax*

The third step was aimed at considering the physical behaviour of the climatic predictors included in the ten models produced at Step 2. Indeed, beside the quantitative performance, it was essential that the climatic predictors physical meaning would be coherent with the known underlying processes. An increasing susceptibility is expected for increasing threshold exceedance frequency for EWIind, EWIper, WD and FT. For snow-related predictors, an increasing susceptibility level is expected in correspondence of high values of SWEep and high negative values of SWEmax. Starting from the

"Visibility" models, "Physically plausible" models were produced meeting two criteria: (i) at least two climatic predictors need to respect their physical meaning in the model; (ii) a maximum of two climatic predictors are allowed with implausible behaviour and would be cast out from the model. If the criteria were not met, the models were discarded, conversely if all the predictors respect their physical meaning, the original "Visibility" model was retained.

The fourth step was aimed at reducing the concurvity among climatic predictors and producing a parsimonious model. According to Wood (2017), concurvity is based on the idea that each smooth term in a model can be decomposed into two parts, one that lies entirely in the space of one or more other terms in the model, and another part that is completely within the term's own space. Concurvity varies from 0, indicating no issues, and 1, indicating total lack of identifiability. A rule of thumb threshold to define high concurvity values is usually set at 0.8 (e.g., Camera et al., 2021). A Principal Component Analysis was carried out in Matlab®, including the grid derived climatic predictors and the snow-melting predictors, to produce a set of new uncorrelated orthogonal principal components (Abdi and Williams, 2010). The calculation of the principal component was carried out on a pixel basis (i.e., for each pixel discretizing the study area the principal components score values are calculated). Then, two "PCA models" were produced by using a number of principal components together explaining at least the 70% of total variance. The models' characteristics are summarized in Table 4.10.

Table 4.10 Model ID and predictors set used in the two "PCA" models produced, one for configuration A and one for configuration B. Note that all the models containing climatic predictors were constructed using the not-penalized morphometric predictors from the TOPO-A and TOPO-B models. The number of principal components included as predictors would depend on the total variance explained in the PCA.

Configuration	***Model ID***	***Morphometric predictors***	***Climatic predictors***
A	PCA ModelA	Not-penalized predictors from TOPO-A	Principal components (PC1, PC2..PCn)
B	PCA ModelB	Not-penalized predictors from TOPO-B	Principal components (PC1 ,PC2..PCn)

Finally, the blind models, the physical plausible models from Step 2 and Step 3, and the PCA models were compared and discussed in terms of performance, climatic predictors' behaviour and role, susceptibility spatial distribution. The output susceptibility maps were reclassified into five susceptibility classes (0.0-0.3 "very low", 0.3-0.5 "low", 0.5-0.7 "medium", 0.7-0.9 "high", 0.9-1.0 "very high"). Some susceptibility variation maps were also prepared in order to further discuss and analyse the different susceptibility spatial patterns.

4.7.2. Results

4.7.2.1 Traditional geo-environmental predictors

Figure 4.31 shows the DEM-derived predictors and the categorized geological predictor.

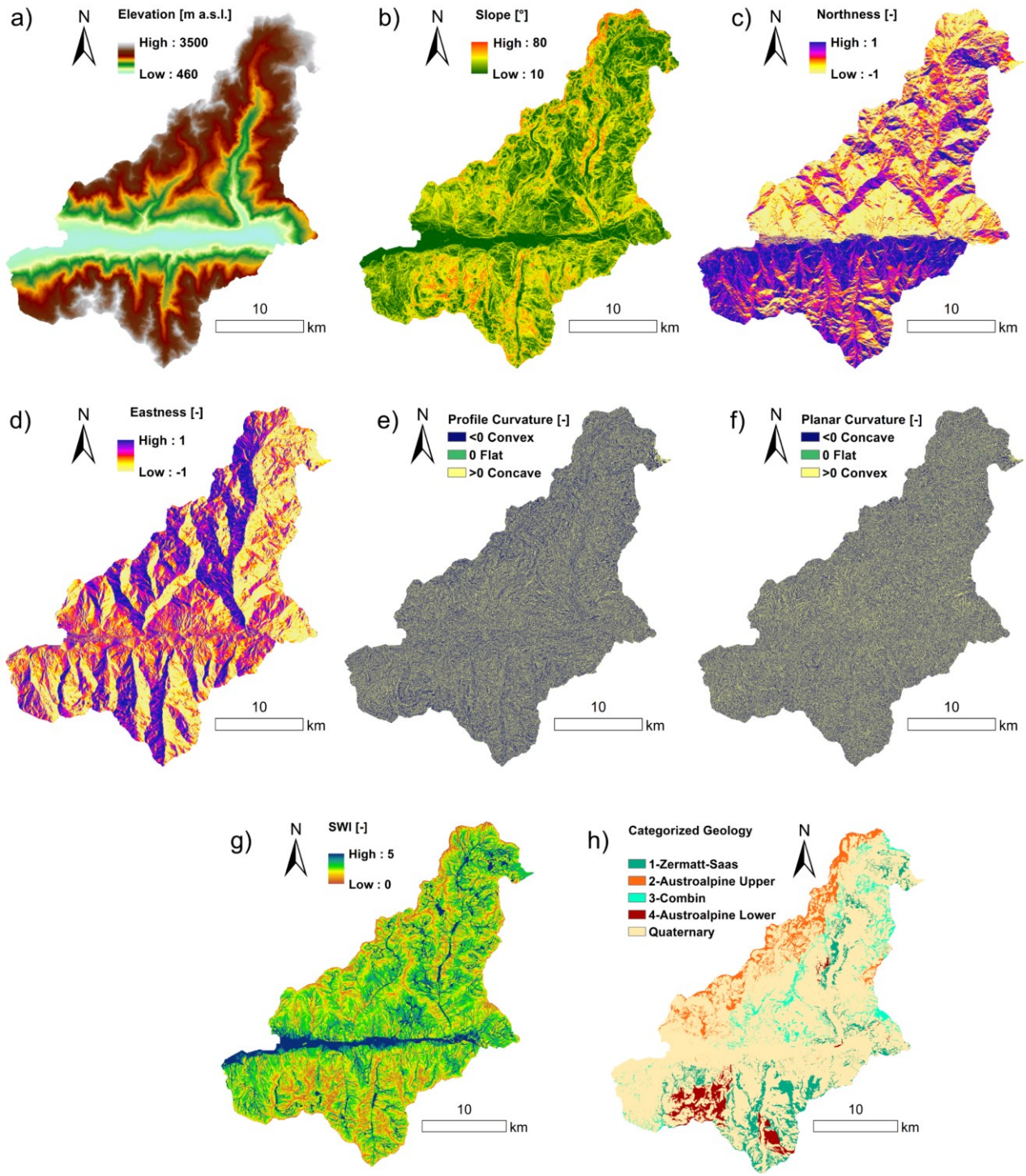


Figure 4.31 maps of the other traditional geo-environmental predictors used as inputs in the susceptibility models, elevation, slope, northness, eastness, profile curvature, planar curvature, SWI (Saga Topographic Wetness Index) and reclassified geology.

4.7.2.2 The "blind" model – Step 1

Variable penalization resulted in the exclusion of the predictors northness, plan curvature and SWI (all characterized by a horizontal CSF) from the Blind-TOPO model (Table 4.11). Therefore, these predictors were excluded from the successive Blind-CLIMATE and Blind-GEO models. Eastness was also frequently penalized in the CV runs but was retained because its CSF on the entire dataset was not completely horizontal.

Table 4.11 Penalization frequency and mDD% for the "Blind" models.

Model ID	**Predictors**	**Penalization Frequency (edf<0.7) [%]**		**mDD%**
		nsCV	sCV	
Blind-TOPO	*Elevation*	0.0	0.0	76.6
	Slope	3.4	7.2	7.2
	Profile curvature	0.2	0.0	5.9
	Plan Curvature	97.6	91.2	4.6
	Northness	99	89.0	4.6
	Eastness	79.4	81.0	0.7
	SWI	86.6	80	0.4
Blind-CLIMATE	*Elevation*	74.8	72.2	10.5
	Slope	0.0	3.2	13.0
	Profile curvature	0.0	0.0	13.3
	Eastness	97.4	91.6	11.5
	EWIind	2.6	0.2	11.5
	WD	0.0	0.0	13.5
	FT	6.2	32.8	8.8
	SWEep	0.4	13.6	8.9
	SWEmax	68.0	68.0	9.0
Blind-GEO	*Elevation*	58.2	34.8	10.2
	Slope	1.4	11.8	11.6
	Profile curvature	0.0	0.0	12.5
	Eastness	98.2	86.2	11.0
	EWIind	9.0	0.0	10.7
	WD	0.0	0.0	15.1
	FT	4.2	27.0	8.9
	SWEep	3.2	6.0	9.7
	SWEmax	76.6	72.2	9.7
	Geology	-	-	0.3

For component smoothing functions (CSF), edf and mDD% for nsCV and sCV of the Blind-CLIMATE model refer to Figure 4.32. In the Blind-CLIMATE model the variables WD, profile curvature, slope, and EWIind were the most important not-penalized variables (Table 4.11). Elevation was frequently penalized in the CV runs, probably due to high concurvity effects with EWIind and FT (both regionalized with the elevation covariate, see Section 4.5.1.2). The CSF for slope and profile curvature, indicating that the modelled rockfall likelihood was highest for high inclined (approximately >50°) and convex slopes, were geomorphologically reasonable and the curve shapes were consistently maintained through

the different nsCV and sCV runs. Eastness, a component of slope aspect, was further penalized when adding the climatic predictors.

The CSF for elevation (Figure 4.32), describing a linear and indirect relationship between elevation and rockfall occurrence, indicated the highest rockfall likelihood corresponded to the lowest altitudes (approximately <1800 m a.s.l.). This is a clear symptom of rockfall overrepresentation close to roads and principal infrastructures, being a data collection effect. Regarding CSF related to the climatic predictors (Figure 4.32), the two only predictors with curves describing a physically plausible behaviour (and consistently maintained through the nsCV and sCV runs) were WD and SWEmax, indicating that the modelled landslide likelihood was highest for the highest WD-related threshold exceedance frequency and in correspondence of areas with the highest cumulated snow melting. The bell shapes of the CSF of SWEep and FT were physically uninterpretable (Figure 4.32), while EWIind had very high dispersion for low threshold exceedance values, which hampers the definition of a clear expected increasing behaviour. The high rockfall likelihood for the highest EWIind threshold exceedance frequency may be a contemporary data collection effect, as the station-based EWIind predictor showed a linear direct relationship with elevation (see Section 4.5.1.2) and a real susceptibility effect

Accounting for the influence of geology (model Blind-GEO), CSF, predictors importance and penalization frequency remained consistent (Table 4.11). The categories Combin ophiolites and Austroalpine lower outliers in the eclogitic facies resulted in the lowest chance of being linked to rockfall susceptibility with odds ratio of 0.5 and 0.58, respectively. The Austroalpine upper outliers in the blueshist facies (odds ratio 1.1), instead had a comparable chance of rockfall likelihood to the Zermatt-Saas unit. However, none of the odds ratios calculated had statistical significance (i.e., p-value >0.05), which, in addition with a very low mDD% of 0.3%, indicates that the geological predictor may have very little influence on rockfalls in the study area. Another possible explanation is the bias in the inventory, which could result in an underrepresentation of rockfalls in specific lithologies, as for the Austroalpine upper outliers, which outcrop almost exclusively at high altitudes in the north-western part of the study area. This was also corroborated by the fact that for the sCV, due to the limited number of rockfalls in the Austroalpine upper and lower outliers (i.e., class II and IV), some cross validation runs did not converge due to absence of training points (Guzzetti et al., 2006) and exacerbated in presence of inventory bias.

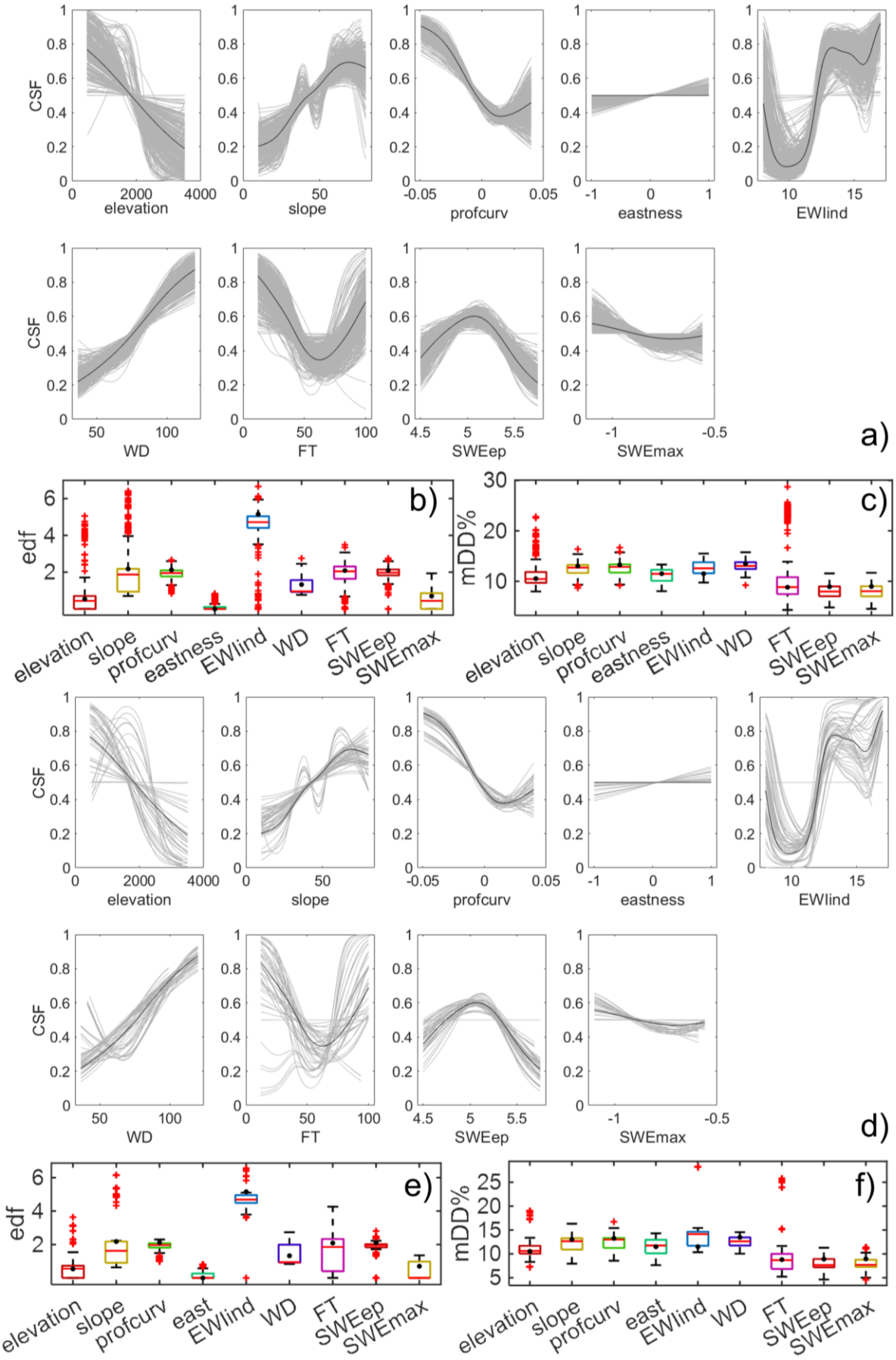


Figure 4.32 Blind-CLIMATE model predictors CSF, edf and mDD% for non-spatial CV (a,b,c) and spatial CV (d,e,f).

The AUROCs indicated a similar and "excellent" (according to Hosmer et al., 2013 guidelines) discrimination capability for all the three "Blind" models (Figure 4.33), with a median AUROC scores of 0.88 (nsCV) and 0.84 (sCV) for the Blind-TOPO model, 0.90 (nsCV) and 0.87 (sCV) for the Blind-CLIMATE model and 0.89 (nsCV) and 0.85 (sCV) for the Blind-GEO model. The climatic predictors, independently from their physical significance, have led to a model performance increase, but to a reduced spatial transferability (i.e., a wider AUROC inter-quantile range for the models including climatic predictors compared to the Blind-TOPO model). The lower performance of the Blind-GEO model respect to the Blind-Climate model confirmed that the geological predictor reduced the discrimination capacity and spatial transferability of the model. Without considering the distorted relationships between some predictors and susceptibility, as well as the presence of a bias in the inventory, similar quantitative performances may induce to trust the "Blind" models. However, it has been demonstrated that this type of models may rely upon non-physically plausible behavior of process-describing climatic predictors and on the influence of data collection procedures (i.e., altitude and, indirectly, vicinity to roads and infrastructures).

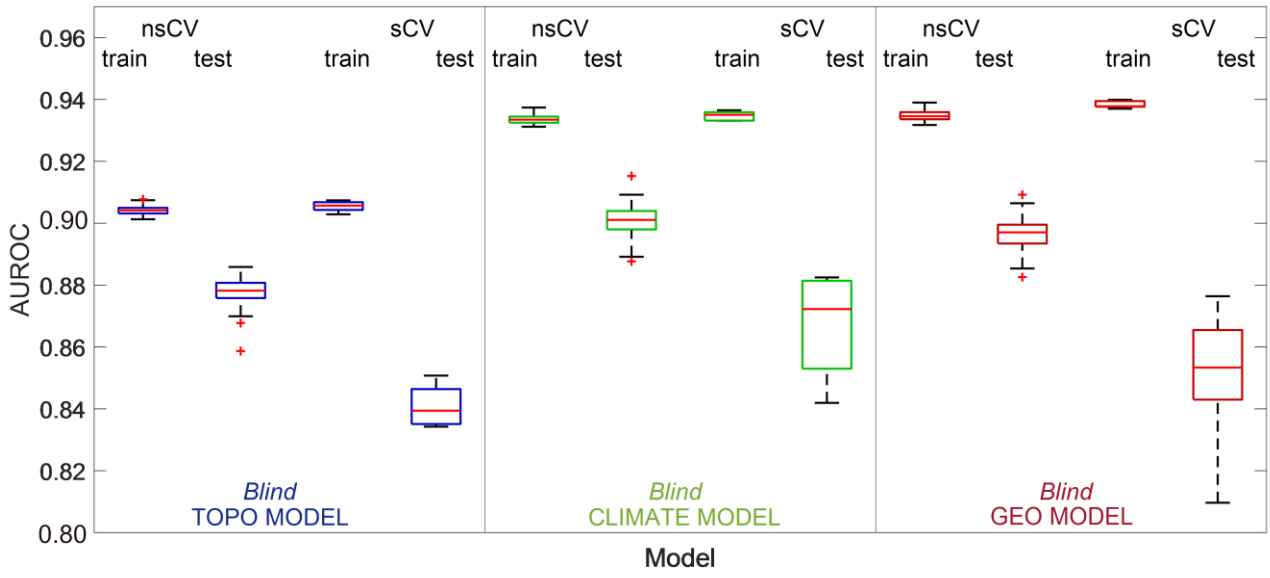


Figure 4.33 Boxplots showing the variability observed during the k-fold non-spatial (nsCV) and spatial (sCV) cross-validation for the BLIND models, regarding the area under the ROC curve (AUROC) values for train and test points.

4.7.2.3 The visibility approach derived models – Step 2

Configuration A and configuration B visibility masks and the correspondent selected negative points are shown in Figure 4.34. They were the results of the bias reduction procedure to compensate the unavailability of rockfall data far from roads and infrastructures. The basic topography-based models (i.e., TOPO models) resulted in the variable penalization and consequent exclusion of the predictors eastness, plan curvature and SWI, both for configuration A and B (Table 4.12).

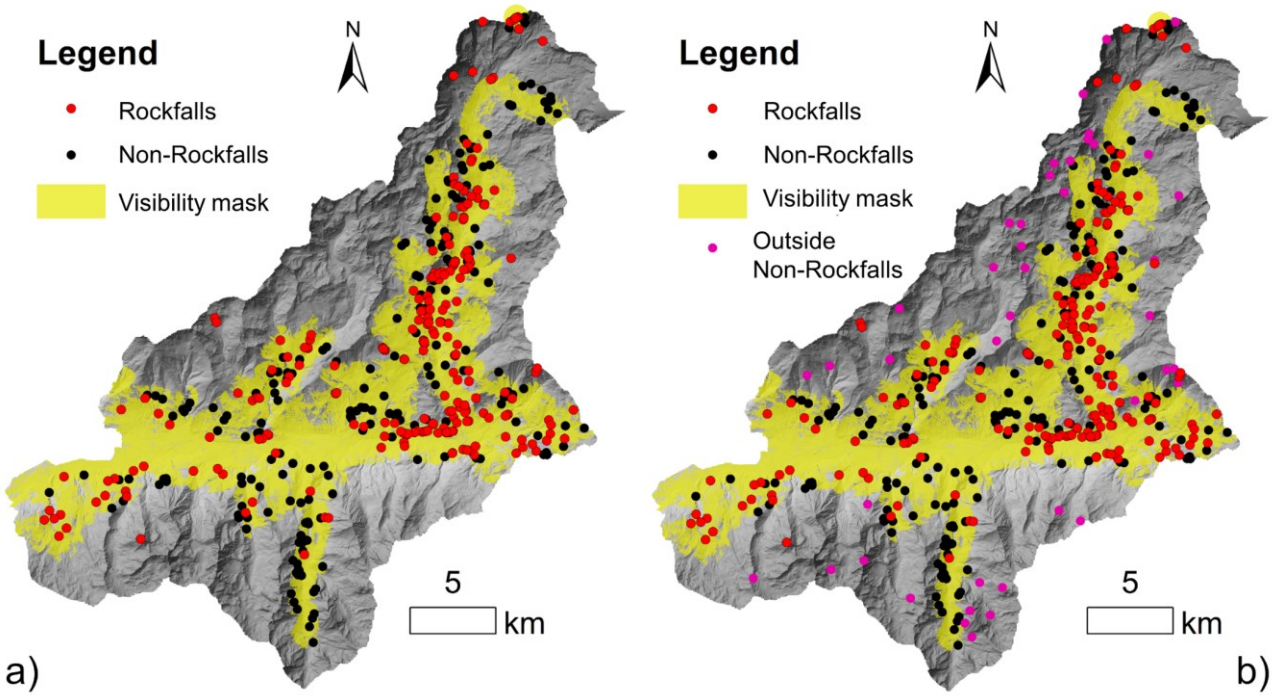


Figure 4.34 Visibility masks and selected non-rockfall points a) Configuration A and b) Configuration B.

Therefore, in order to produce parsimonious models, these predictors were excluded from the successive climate models (i.e., a total of eight models). The remaining morphometric variables (elevation, slope, profile curvature and northness) resulted to have almost identical CSF shapes, penalization frequency and deviance explained in the TOPO models of configuration A and B, also essentially maintained in all the successive eight climatic models tested. Their representative CSF, edf and mDD% coming from non-spatial and spatial CV, are visible in Figure 4.37, Figure 4.38 in Figure 4.39 and are summarized in Table 4.12. These results highlighted the consistency of the behaviour of the important morphometric predictors, which was altered neither by the addition of climatic predictors nor by a different selection of non-landslide points. The CSF for slope and profile curvature indicated that the modelled rockfall likelihood was highest for high inclined (approximately >50°) and convex slopes. The CSF of northness was represented by a linear, gently inclined function (in some models very close to penalization) indicating that the modelled rockfall likelihood was only slightly higher for south facing slopes than for north facing slopes. Due to solar radiation, the slopes facing south are the most exposed to diurnal cycles of temperature and should experience the highest rates of wet and dry cycles and snowmelt in winter. The CSF of elevation had a wavy trend with three peaks at low (approximately 500 m a.s.l.), medium (approximately 1500 m a.s.l.) and high (approximately 3000 m a.s.l) altitudes. The low elevation peak is the highest. This may indicate either that the bias effect was not totally removed through the visibility approach or that this trend actually reflects the real spatial distribution of productive rock walls in the study area.

Table 4.12 Penalization frequency and mDD% for morphometric predictors in all the models for configuration A and B.

Model		Morphometric predictor	Penalization Frequency (edf<0.7) [%]		mDD%
			nsCV	sCV	
A	TOPO-A	*Elevation*	0.0	0.0	42.0
		Slope	0.0	0.0	16.7
		Profile Curvature	12.4	26.6	15.0
		Plan curvature	99.6	100.0	3.7
		Northness	0.0	0.0	9.3
		Eastness	94.4	89.2	6.8
		SWI	97.6	97.6	6.4
	ST Model1A	*Elevation*	0.0	0.0	25.7
		Slope	0.6	0.8	16.8
		Profile Curvature	0.0	0.0	17.4
		Northness	88	78.6	8.7
	ST Model2A	*Elevation*	0.0	0.0	30.0
		Slope	0.4	0.2	15.6
		Profile Curvature	0.4	0.0	16.0
		Northness	88.6	82.8	10.0
	GR Model1A	*Elevation*	0.0	0.0	28.7
		Slope	0.0	13	13.8
		Profile Curvature	0.0	0.0	12.4
		Northness	57.8	78.8	10.3
	GR Model2A	*Elevation*	0.0	0.0	27.4
		Slope	0.0	0.0	15.2
		Profile Curvature	0.0	0.0	13.5
		Northness	60.6	76.2	10.3
B	TOPO-B	*Elevation*	0.0	0.0	41.5
		Slope	0.0	0.0	19.4
		Profile Curvature	0.0	0.0	14.0
		Plan curvature	100.0	97.2	0.55
		Northness	23.2	40	13.0
		Eastness	100.0	100.0	11.2
		SWI	99.6	99.0	0.55
	ST Model1B	*Elevation*	4.4	0.2	17.6
		Slope	0.0	0.0	15.7
		Profile Curvature	0.0	0.0	14.8
		Northness	75.8	68.2	15.0
	ST Model2B	*Elevation*	0.0	0.0	17.8
		Slope	0.0	0.0	15.8
		Profile Curvature	0.0	0.0	15.0
		Northness	78.6	80.6	15.2
	GR Model1B	*Elevation*	0.0	0.0	22.8
		Slope	0.0	1.2	14.9
		Profile Curvature	0.0	0.0	12.4
		Northness	41	54.2	12.6
	GR Model2B	*Elevation*	0.2	0.0	19.4
		Slope	0.0	0.0	14.3
		Profile Curvature	0.0	0.0	12.4
		Northness	33.2	49.6	12.5

Regarding the climatic predictors, being threshold exceedances, a susceptibility increase is expected as their value increases. Therefore, only nearly monotonous increasing smoothing functions would be considered as an acceptable physically plausible behaviour. The same applies for the number of snowmelt events (SWEep), while for SWEmax a monotonous decreasing function is expected as the more negative is its value, the higher is the water input coming from snow melting processes. The behaviour, penalization frequency and mDD% of the eight models (station-based or grid-based, configuration A or B) are reported in Table 4.13 and the main results are summarized below:

(i) EWIper had always a physically implausible behaviour, with a bell shape in ST Model 2A and a monotonous decreasing function for all the other models in which it was included (GR Model 2A, ST Model 2B, GR Model 2B).

(ii) EWIind had a physically plausible behaviour in configuration A, both for the station-based and grid-based models (ST Model1A and GR Model 1A) and in the grid-based configuration B model (GR Model 1B). Conversely it resulted in a physically implausible behaviour in the station-based model for configuration B (ST Model 1B). Moreover, it was the most important climatic predictor in models GR Model 1A and GR Model 1B, with an mDD% > 10% and the second most important climatic predictor for ST Model 1A, with an mDD% around 9%.

(iii) WD always had a physically plausible behaviour for all the four station-based models (ST Model 1A, ST Model 2A, ST Model 1B and ST Model 2B), showing a linear steep CSF and ranking second in terms of importance with an mDD% ranging from 6.6% to 9.2%. The exception was ST Model 1A, in which it was the most important climatic predictor with an mDD% of 10.9%. Conversely, it was penalized in all the grid-based models (GR Model 1A, GR Model 2A, GR Model 1B and GR Model 2B).

(iv) FT resulted to be physically significant only for the station-based models in configuration A (ST Model1A and ST Model 2A), with a linear increasing CSF and an mDD% of about 3%. This predictor was not physically plausible in GR Model 1B and it was penalized in all the other models (ST Model 1B, ST Model 2B, GR Model 1A, GR Model 2A and GR Model 2B).

(v) SWEep was penalized in ST Model 2A and had a physically implausible behaviour in all other models.

(vi) SWEmax showed a physically plausible behaviour for all the models of configuration A (ST Model 1A, ST Model 2A, GR Model 1A and GR Model 2A), and for GR Model 1B, with an mDD% ranging approximately between 3% and 5%. It was penalized in the station-based models of configuration B (ST Model 1B and ST Model 2B) and not physically plausible in GR Model 2B.

Table 4.13 Physical significance (yes/no) penalization frequency and mDD% for climatic predictors in the all the tested models including the climatic component (both configuration A and B models).

Model		Source of Climate predictors	Climate predictor	Physical plausibility	Penalization Frequency(edf<0.7) %		mDD%
					nsCV	sCV	
A	ST Model1A	Station	*EWIind*	**yes**	32.6	55.0	9.1
			WD	**yes**	0.0	9.0	10.9
			FT	**yes**	78.2	51.8	2.7
			SWEep	**-**	55.8	51.6	3.7
			SWEmax	**yes**	36.4	26.0	4.8
	ST Model2A	Station	*EWIper*	**no**	70.4	31.4	9.9
			WD	**yes**	0.0	7.4	9.2
			FT	**yes**	59.8	35.5	2.8
			SWEep	**no**	22.6	19.6	3.6
			SWEmax	**yes**	64.4	66.6	3.1
	GR Model1A	Gridded	*EWIind*	**yes**	9.6	21.0	10.2
			WD	**-**	92.8	71.6	6.6
			FT	**-**	99.2	95.0	6.4
			SWEep	**no**	18.6	20.6	6.4
			SWEmax	**yes**	9.0	11.4	5.1
	GR Model2A	Gridded	*EWIper*	**no**	67.0	60.4	9.9
			WD	**-**	83.8	71.6	6.3
			FT	**-**	97.2	94.4	6.0
			SWEep	**no**	18.4	31.8	5.9
			SWEmax	**yes**	1.6	13.8	5.4
B	ST Model1B	Station	*EWIind*	**no**	28.6	43.2	16.5
			WD	**yes**	0.6	9.4	9.2
			FT	**-**	95.0	88.6	4.0
			SWEep	**no**	14.6	18.6	4.2
			SWEmax	**-**	83.6	57.8	3.1
	ST Model2B	Station	*EWIper*	**no**	9.4	13.2	16.7
			WD	**yes**	0.0	10.6	9.2
			FT	**-**	88.6	52.4	3.7
			SWEep	**no**	0.6	10.4	3.9
			SWEmax	**-**	95.6	88.6	2.8
	GR Model1B	Gridded	*EWIind*	**yes**	0.2	0.4	13.5
			WD	**-**	82.8	86	6.4
			FT	**no**	86.8	52.6	6.4
			SWEep	**no**	1.8	2.2	6.8
			SWEmax	**yes**	53.8	44.0	4.3
	GR Model2B	Gridded	*EWIper*	**no**	79.6	73.6	12.3
			WD	**-**	93.4	93.2	7.7
			FT	**-**	76.2	59.0	7.7
			SWEep	**no**	0.4	0.0	8.0
			SWEmax	**no**	11.6	31.4	5.7

From these results, the general outcome was that a slight change in the negative selection configuration and the alternative use of the station-based and grid-based climatic predictors, with their own specific spatial patterns, could moderately affect the model response and the interrelationships between climatic

predictors. However, some trademarks were still recognizable, in terms of predictors importance ranking and smoothing function shapes.

Looking exclusively at model performance, the best mode for Configuration A (Figure 4.35) was ST Model 1A, with a mean AUROC increment respect to the TOPO-A model of approximately 3% both for the nsCV and sCV. In terms of nsCV all the models including climatic predictors showed a good discrimination capability with mean AUROCs between 0.75 and 0.80 and a better performance than the TOPO-A model. In general, the two station-based models (ST Model 1A and ST Model 2A) exhibited higher discrimination capabilities than the corresponding grid-based models (i.e., GR Model 1A and GR Model 2A).

Also, the models including the EWIind variable (ST Model 1A and GR Model 1A) showed better performance than the models including the alternative EWIper predictor (ST Model 2A and GR Model 2A). In terms of sCV, the discrimination capability became only acceptable, with mean AUROCs between 0.6 and 0.7 for all models. The only model with a better performance than the TOPO-A model was again the ST Model 1A. However, it also resulted to have the highest AUROC associated IQR (i.e., the lowest spatial transferability). Moreover, in the case of sCV, the GR Model 1A had a slightly lower performance than the grid-based model including EWIper (GR Model 2A), but it showed the lowest IQR (i.e., the highest spatial transferability).

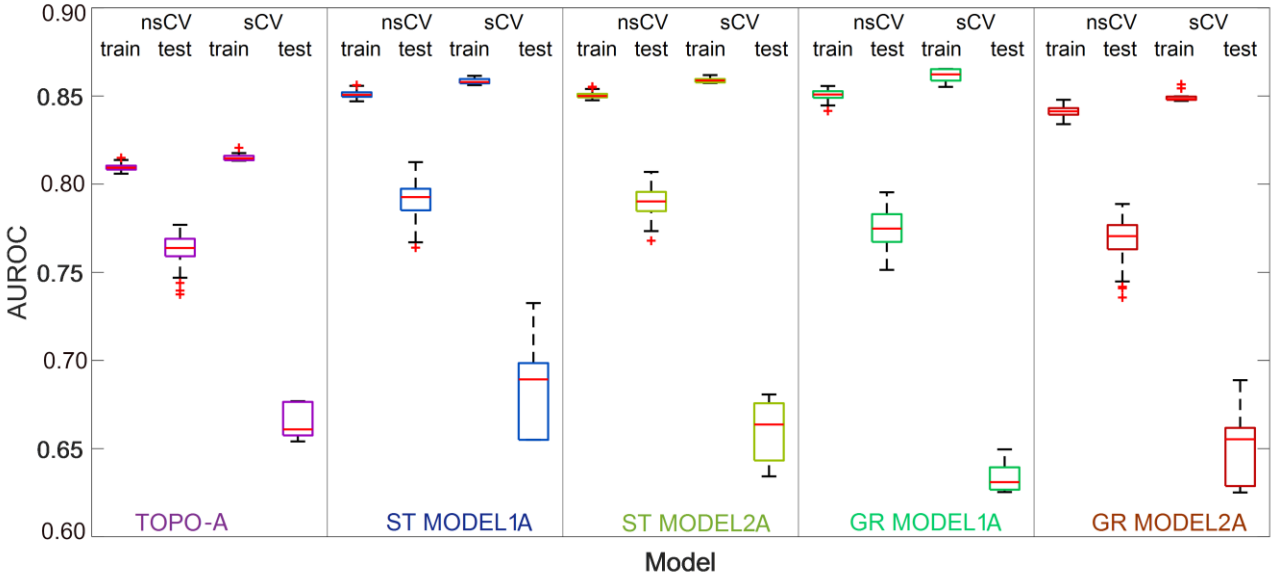


Figure 4.35 Boxplots showing the variability observed during the k-fold non-spatial (nsCV) and spatial (sCV) cross-validation of the Configuration A related models, regarding the area under the ROC curve (AUROC) values for train and test points.

For Configuration B (Figure 4.36), in terms of nsCV all the models including climatic predictors had AUROC between 0.75 and 0.8 and had better discrimination capabilities than the TOPO-B model. ST Model 2B had the highest nsCV mean AUROC (0.78), followed by the equally performing ST Model 1B and GR Model 1B (mean AUROC of 0.77) and GR Model 2B (mean AUROC of 0.76). In terms of sCV, the discrimination capability of all models became only acceptable (mean AUROC between 0.60 and 0.70). The only model with a better performance than TOPO-B was ST Model 1B, which however showed the highest IQR (lowest spatial transferability), similarly to what happened for configuration A. As a common thread between the two configurations, the inclusion of climatic predictors slightly increased the model performance for nsCV, while it did not enhance the mean AUROC related to the sCV and decreased the spatial transferability of models. Moreover, if compared with the "blind" models' performances, an AUROC reduction was obtained as a fee for the application of a bias reduction strategy.

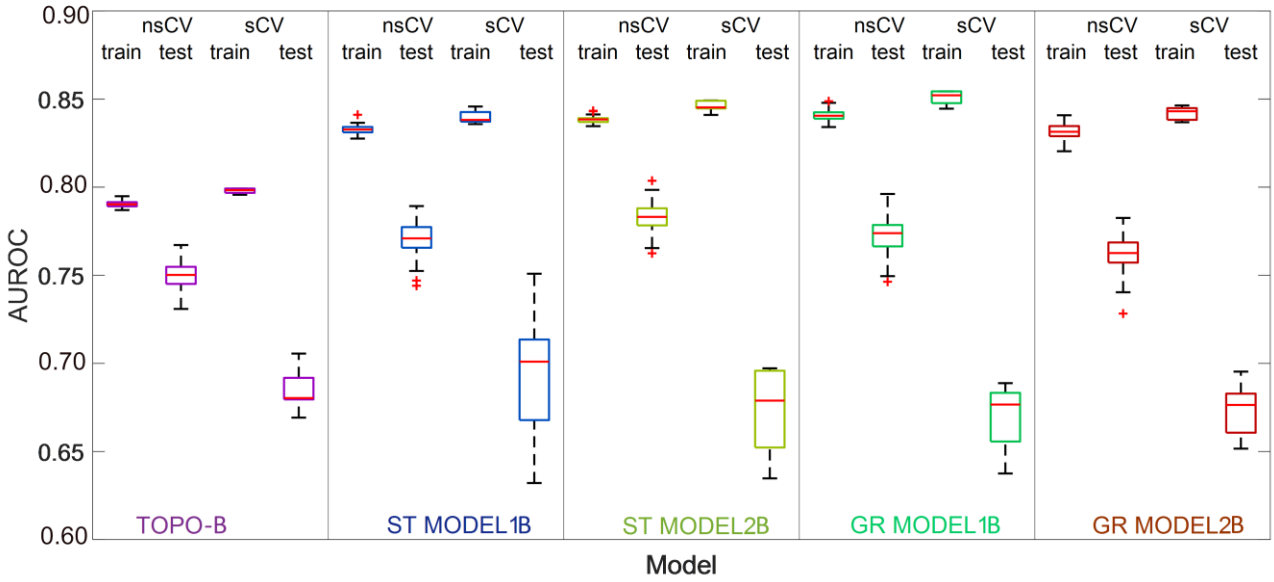


Figure 4.36 Boxplots showing the variability observed during the k-fold non-spatial (nsCV) and spatial (sCV) cross-validation of the Configuration B related models, regarding the area under the ROC curve (AUROC) values for train and test points.

4.7.2.4 Physically plausible models – Step 3

All models with EWIper (ST Model 2A, GR Model 2A, ST Model 2B, GR Model 2B) were discarded as they presented more than two physically implausible climatic predictors and less than two physically plausible ones (criteria defined in Section 4.7.1.3). The exception was ST Model 2A, but also in this case the model was discarded because its counterpart with EWIind (ST Model 1A) performed better (all predictors plausible and one predictor penalized). ST Model 1B was also discarded, as only WD was physically significant, and the other climatic predictors were penalized or not physically plausible. The remaining models (ST Model 1A, GR Model 1A and GR Model 1B) were potentially optimal, as they

maximised, for the associated visibility configuration, the number of included climatic predictors concurring in a physically plausible way in defining susceptibility. In ST Model 1A, all climatic predictors were physically plausible, except for SWEep, which was statistically penalized. Therefore, the model was maintained as it was. In GR Model 1A and GR Model 1B, the physically implausible predictors (SWEep and SWEep plus FT, respectively) were casted out from the models. Consequently, two reduced grid-based models (GR-RedA and GR-RedB) were produced.

For the ST Model 1A (Figure 4.37 and Table 4.14), the importance and CSF of morphometric predictors (elevation, slope, profile curvature and northness) were already discussed and presented in the previous Section 4.7.2.3. Regarding climatic predictors, WD resulted to be the most important not-penalized variable, followed by EWIind, SWEmax and FT. The medium-high penalization frequency in the CV runs and the low mDD% of FT may be imputed to its superimposition, and associated concurvity, with elevation. All the CSF related to the climatic predictors were represented by approximately monotonic increasing (decreasing for SWEmax) functions with different steepness, reflecting predictors importance (e.g., the CSF of WD was the steepest). These behaviours indicated that the modelled rockfall likelihood was highest for the highest threshold exceedance frequencies for WD, EWIper and FT and in correspondence of areas with the highest cumulated snow melting (SWEmax). These relationships were consistently maintained across the majority of the nsCV and sCV runs, except for some bell shape CSF observed in some cross-validation runs for EWIind and FT.

Table 4.14 Penalization frequency and mDD% for ST Model1A

Model	**Morphometric predictor**	**Penalization Frequency (edf<0.7) [%]**		**mDD%**
		nsCV	sCV	
ST Model1A	*Elevation*	0.0	0.0	25.7
	Slope	0.6	0.8	16.8
	Profile Curvature	0.0	0.0	17.4
	Northness	88.0	78.6	8.7
	EWIind	32.6	55.0	9.1
	WD	0.0	9.0	10.9
	FT	78.2	51.8	2.7
	SWEep	55.8	51.6	3.7
	SWEmax	36.4	26.0	4.8

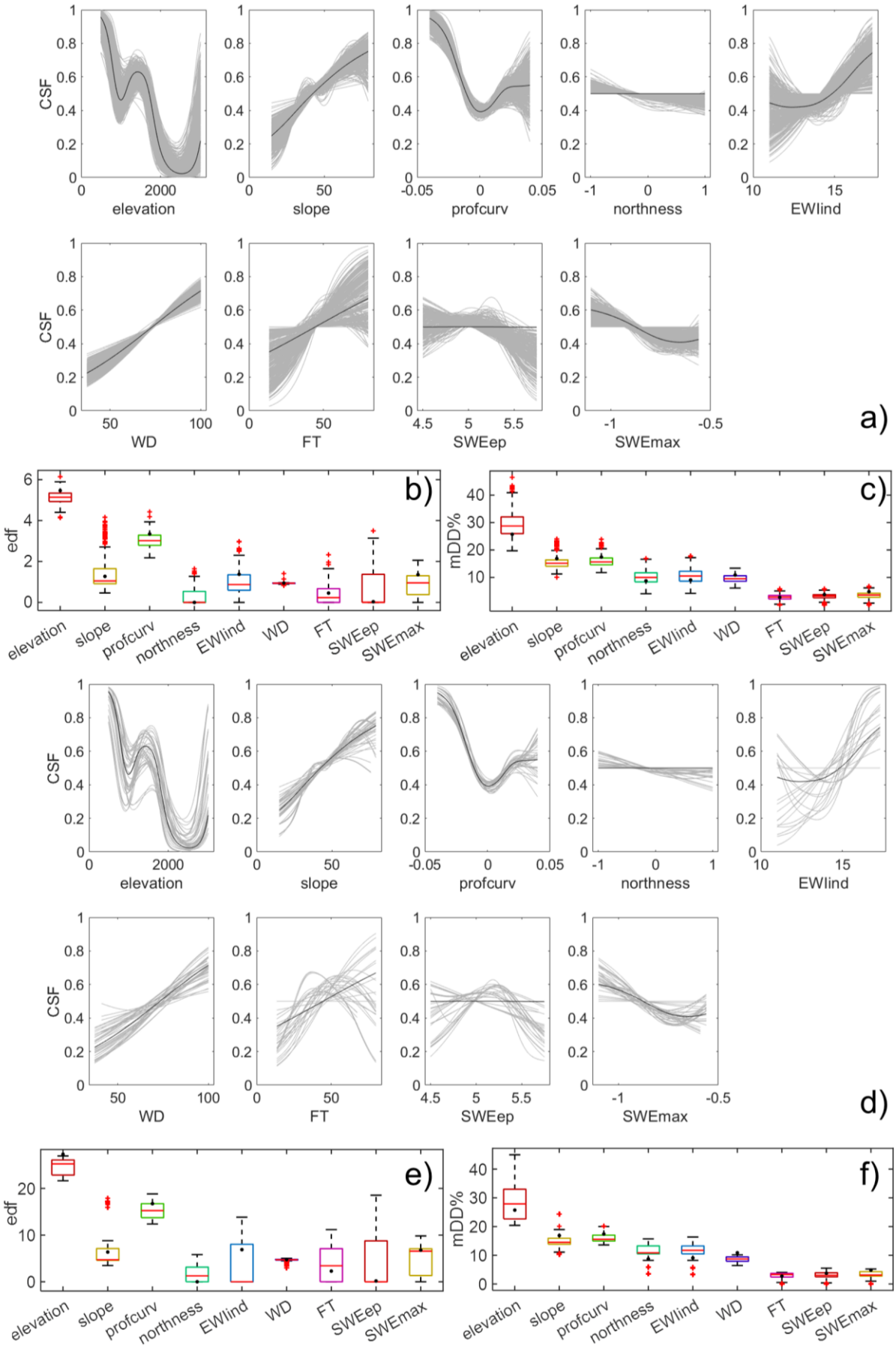


Figure 4.37 ST-Model1A model predictors CSF, edf and mDD% for non-spatial CV (a,b,c) and spatial CV (d,e,f).

For the GR-RedA model (Figure 4.38 and Table 4.15), the importance and CSF of morphometric predictors (elevation, slope, profile curvature and northness) were consistently maintained approximately equal to the ones discussed in the previous Section 4.7.2.3. Despite the elimination of the physically implausible predictor SWEep, these results confirmed the robustness and invariance of morphometric predictors role across models. Among climatic predictors, EWIind resulted to be the most important not-penalized variable, followed by SWEmax and WD. FT remained penalized both on the entire dataset (i.e., CSF as a horizontal line) and in CV runs. As for ST Model 1A, this may be imputed to concurvity effects between FT and elevation. The CSF related to EWIind showed a nonlinear behaviour. At low values (between 14 and 15), the CSF presented a near-sigmoidal shape, probably due to the high dispersion of the data for low threshold exceedance values. For threshold exceedance frequencies higher than 15, the behaviour became nearly linear and increasing. Overall, an increasing trend is clearly visible, despite the initial part of the curve showed CSF coefficients below 0.5 (i.e., indicating a negative contribution to rockfall susceptibility). The CSF of SWEmax and WD were both represented by monotonous, almost linear functions coherent with the variables physical meaning. The above discussed relationships were consistently maintained across the majority of the nsCV and sCV runs.

Table 4.15 Penalization frequency and mDD% for GR-RedA Model

Model ID	**Morphometric predictor**	**Penalization Frequency (edf<0.7) [%]**		**mDD%**
		nsCV	sCV	
GR-RedA	*Elevation*	0.0	0.0	33.7
	Slope	0.0	0.0	16.7
	Profile Curvature	0.0	0.0	14.3
	Northness	71.4	79.0	9.6
	EWIind	8.4	21.0	10.0
	WD	67.0	58.2	4.8
	FT	97.8	95.2	4.3
	SWEmax	10.0	20.0	6.5

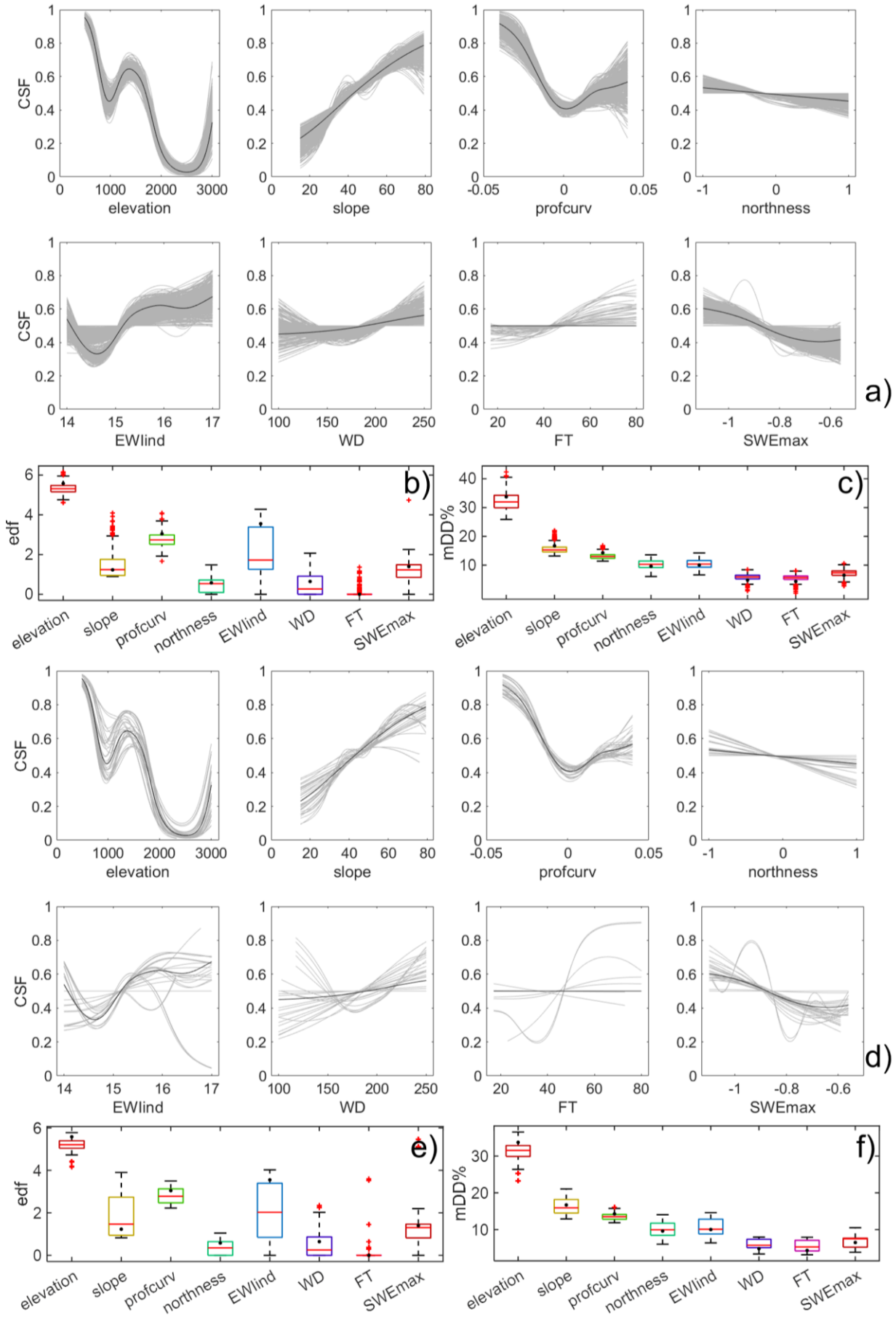


Figure 4.38 GR-RedA model predictors CSF, edf and mDD% for non-spatial CV (a,b,c) and spatial CV (d,e,f).

For the GR-RedB model (Table 4.16 and Figure 4.39), the importance and CSF of morphometric predictors (elevation, slope, profile curvature and northness) were consistently maintained approximately equal to the ones discussed in the previous Section 4.7.2.3. Despite the elimination of the physically implausible predictor SWEep and FT, these results confirmed once again the robustness and invariance of morphometric predictors role across models. Among climatic predictors, EWIind resulted to be the most important not-penalized variable, followed by SWEmax. WD remained penalized both on the entire dataset (i.e., CSF as a horizontal line) and in CV runs. Similarly to the previous GR-RedA model, the CSF related to EWIind showed a nonlinear behaviour. At low values (between 14 and 15), the CSF presented a near-sigmoidal shape. For threshold exceedance frequencies higher than 15, the behaviour became nearly linear and increasing. Also, the CSF appeared steeper than in the GR-RedA model. The CSF of SWEmax appeared mildly decreasing along an almost linear curve, coherent with the variable physical meaning (i.e., the highest rockfall likelihood for the highest cumulated snow melting). The above discussed relationships were consistently maintained across the majority of the nsCV and sCV runs.

Table 4.16 Penalization frequency and mDD% for GR-RedB Model

Model ID	**Morphometric predictor**	**Penalization Frequency (edf<0.7) [%]**		**mDD%**
		nsCV	sCV	
GR-RedB	*Elevation*	0.0	0.0	32.9
	Slope	0.0	0.0	18.4
	Profile Curvature	0.0	0.0	13.5
	Northness	63.8	70.0	12.6
	EWIind	0.0	0.0	15.4
	WD	94.6	80.0	3.3
	SWEmax	64.0	52.6	3.8

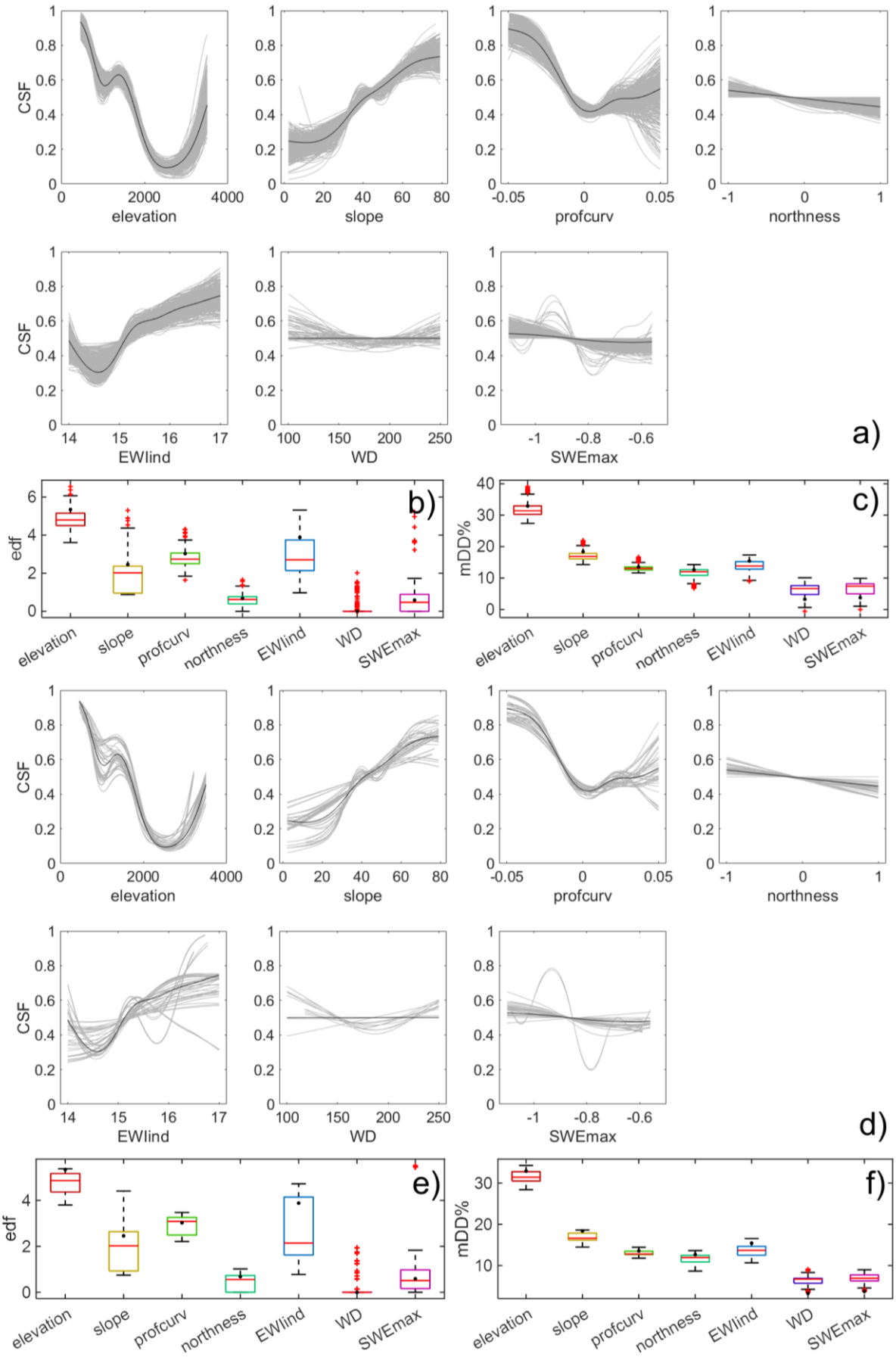


Figure 4.39 GR-RedB model predictors CSF, edf and mDD% for non-spatial CV (a,b,c) and spatial CV (d,e,f).

The quantitative performance, in terms of non-spatial cross validation, was slightly lower for the two reduced grid-based model than for ST Model 1A (mean AUROC 0.79), with a mean AUROC of 0.77 and 0.76 for GR-RedA and GR-RedB, respectively. In terms of spatial cross validation, the performance gap was wider: a mean AUROC of 0.62 for GR-RedA model and 0.65 for GR-RedB, against a mean AUROC of 0.69 for ST Model1A (Figure 4.40). Using the 42 rockfalls excluded from Configuration A models as an independent holdout test set, the AUROC was 0.80 for ST-Model1A and 0.82 for GR-RedA model.

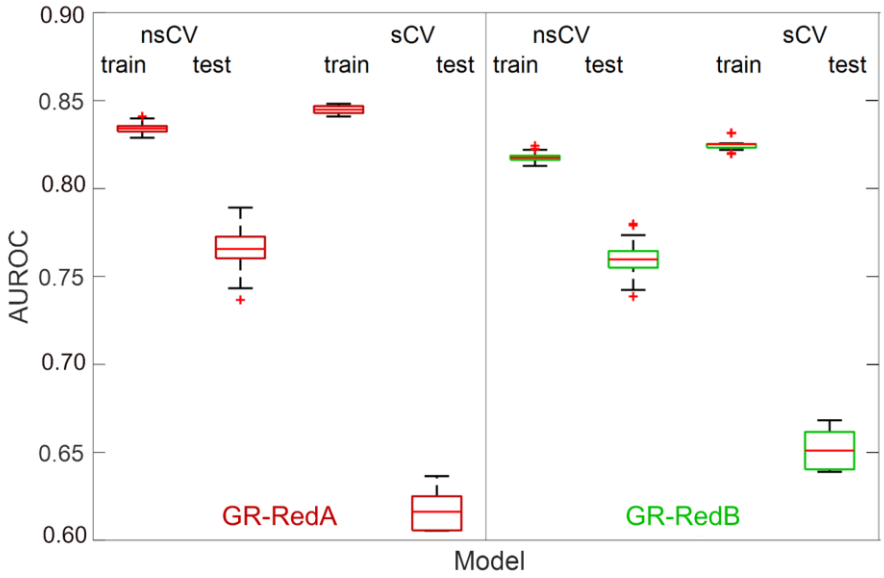


Figure 4.40 Boxplots showing the variability observed during the k-fold non-spatial and spatial cross- validation of the GR-RedA and GR-RedB models, regarding the area under the ROC curve (AUROC) values for training and testing landslide points.

Although the statistically and physically guided selection of significant climatic predictors, concurvity resulted to be high and could cause unstable estimates (Wood, 2017). The calculated concurvity from *mgcv* R package among predictors is showed in Table 4.17.

Table 4.17 Calculated concurvity for each predictor in models ST-Model1A, GR-RedA and GR-RedB.

Predictor	**Concurvity**		
	ST-Model1A	GR-RedA	GR-RedB
Elevation	0.97	0.88	0.66
Slope	0.41	0.22	0.18
Profile Curvature	0.22	0.15	0.11
Northness	0.47	0.41	0.34
EWInd	0.93	0.70	0.56
WD	0.88	0.77	0.70
FT	0.95	0.85	-
SWEep	0.80 (penalized)	-	-
SWEmax	0.78	0.52	0.44

An evident concurvity issue between elevation and the FT climatic predictor was recorded in all the three physically plausible models. It was not unexpected, as temperature variations are strictly coupled with altitudinal variations. In ST Model1A, high concurvity values were also recorded for EWIper and WD and SWEmax; the highest concurvity in the model containing the station-based EWIper may be partially due to the regionalization step of the predictor, which included elevation as a covariate. As a proof, concurvity values of these predictors were only medium-high for the grid derived models (GR-RedA and GR-RedB). The presence of concurvity is however physically reasonable, as climate processes are interconnected. As an example, EWIper and WD are both related to the precipitation regime. Also, wet and dry episodes frequency is linked to the thermal regime, which may control the frequency and duration of storms. EWIper contains a water input from snow, thus it may be linked to SWEmax and thermal fluctuations. Nonetheless, the application of shrinkage for variable selection should allow to penalize the portion of the smoothing function involved in multi-concurvity issues (Figueiras et al., 2005; Laceby et al., 2016; Bagalwa et al., 2021). This is possibly the main reason for the high penalization frequency of FT predictors in the majority of the models.

4.7.2.5 PCA models – Step 4

A further reduction of concurvity, at least amongst climatic predictors, was obtained by means of a Principal Component Analysis (PCA) between the five climatic predictors. Two principal components PC1 and PC2, together explaining 74.1% of predictors variance (44.5% and 29.6%, respectively), were used as predictors for the susceptibility models. The first principal component PC1 explained the climatic predictors FT, EWIind and WD with a direct relationship (Figure 4.41), meaning that high and positive PC1 values corresponded to high threshold exceedance frequencies. For this reason, the CSF of PC1 in the rockfall susceptibility model is expected to be a monotonic increasing function.

The second principal component PC2 explained the snow-related predictors SWEmax and SWEep with a direct relationship (Figure 4.41). This means that high positive values of PC2 corresponded to a high number of snow melting events (i.e., high values of SWEep), but to low cumulated value of snow melting (as SWEmax represent a depletion of SWE storage, the more negative are its values, the more cumulated snowmelt). This discrepancy means that the expected CSF function for PC2 should be a semi-linear monotonic increasing function to explain in a physically plausible way SWEep, while should be a semi-linear monotonic decreasing function to explain in a physically plausible way SWEmax. Considering that SWEmax is almost parallel to PC2 axis (Figure 4.41), the effect of this predictor on the definition of PC2 related CSF is expected to be stronger than SWEep.

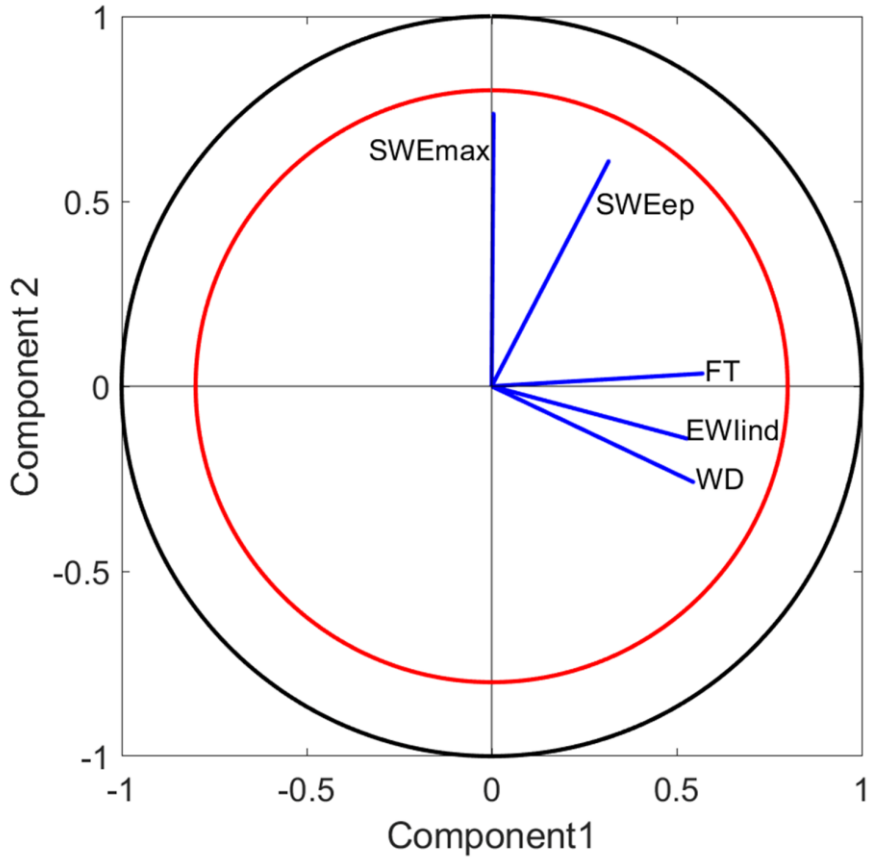


Figure 4.41 Diagram of the principal components PC1 (x-axis) and PC2 (y-axis). Blue vectors represent the climatic predictors; their length describes variable importance in defining principal components (the longer the more important) and their direction defines whether the relationship with the components is positive or negative. For easily representation purposes, the components and variables scores were normalized between -1 and 1: the inner red circle represents a relationship of ± 0.8 and the outer black circle a relationship of ± 1.

For both PCA Model A and PCA Model B, the importance and CSF of morphometric predictors (elevation, slope, profile curvature and northness) were consistently maintained similarly to the previously discussed models. Regarding the principal components, they showed the expected behaviour, almost identical for PCA Model A and PCA Model B (Figure 4.42and Figure 4.43, respectively). For low values (< 0.0), the CSF of PC1 showed a near-sigmoidal shape, probably inherited from EWIind, while it showed an almost linear increasing behaviour for positive values. The CSF of PC2 showed a linear monotonic decreasing function, physically coherent with SWEmax, which had the strongest signal on the component definition. These relationships were consistently maintained across the majority of the nsCV and sCV runs. Predictors importance and penalization frequency for PCA ModelA and PCA ModelB are reported in Table 4.18. The advantage of running the model with principal components instead of the original climatic predictors was that not only the concurvity amongst climatic predictors was consistently reduced, as the principal components are orthogonal thus completely decorrelated, but also concurvity with elevation was decreased below the 0.8 threshold (Table 4.18).

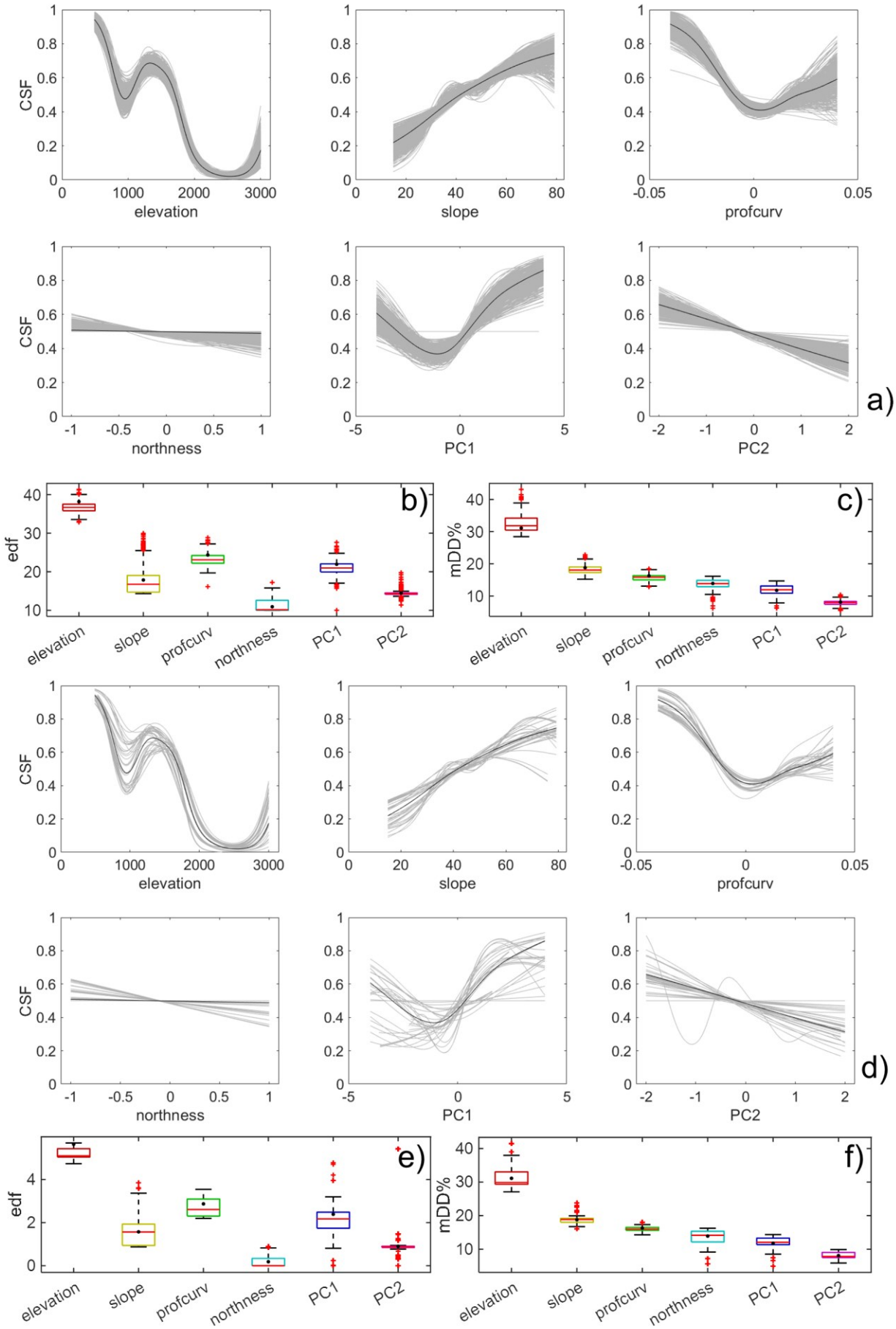


Figure 4.42 PCA ModelA predictors CSF, edf and mDD% for non-spatial CV (a,b,c) and spatial CV (d,e,f).

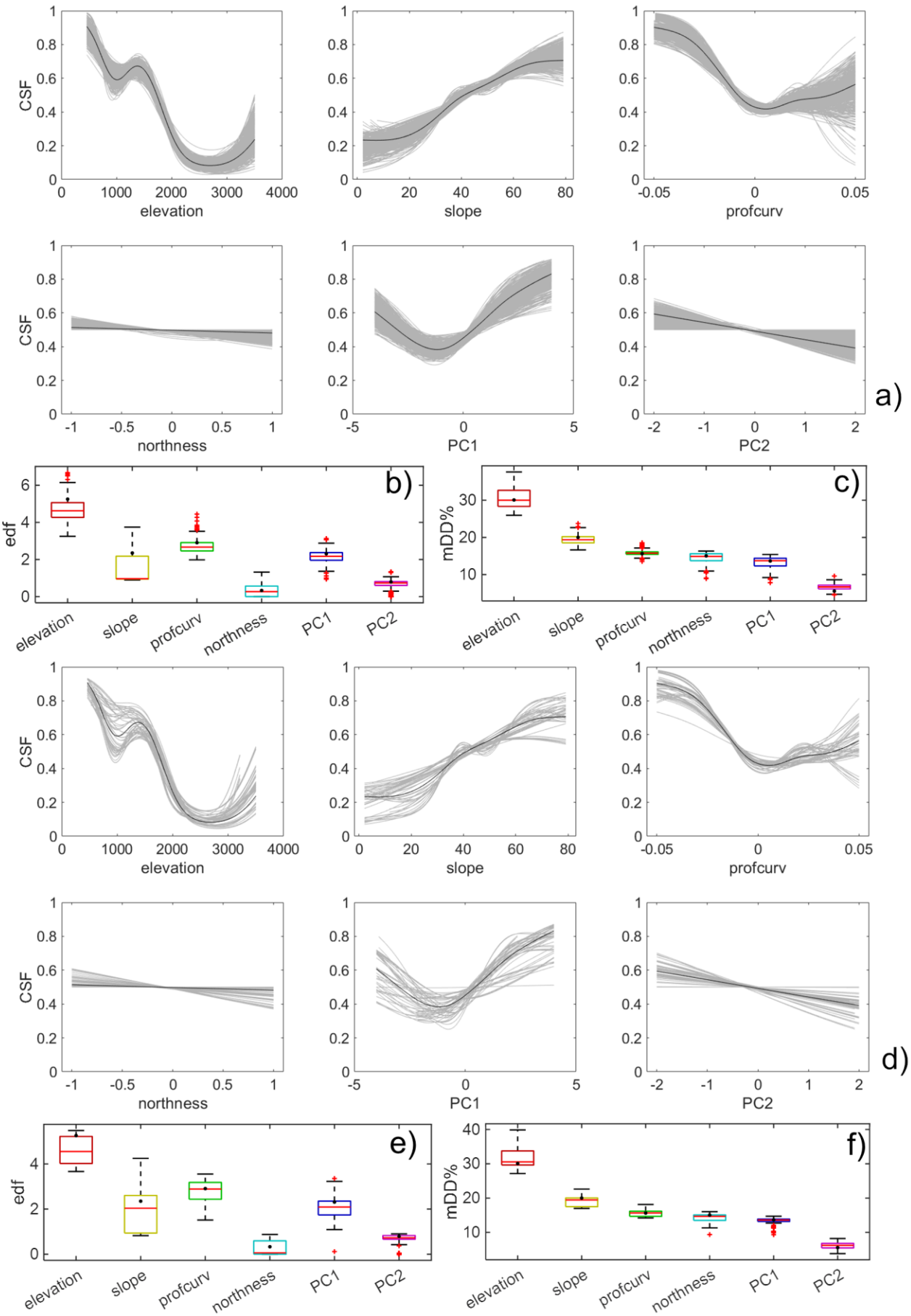


Figure 4.43 PCA ModelB predictors CSF, edf and mDD% for non-spatial CV (a,b,c) and spatial CV (d,e,f).

Table 4.18 Calculated concurvity, penalization frequency and mDD% for each predictor in PCA ModelA and PCA ModelB.

Model ID	Predictor	Concurvity	Penalization Frequency (edf<0.7) [%]		mDD%
			nsCV	sCV	
PCA ModelA	*Elevation*	0.75	0.0	0.0	31.1
	Slope	0.19	0.0	0.0	18.8
	Profile Curvature	0.13	0.0	0.0	16.3
	Northness	0.30	90.8	79.8	13.9
	PC1	0.73	0.2	6.0	11.8
	PC2	0.40	4.2	17.2	8.1
PCA ModelB	*Elevation*	0.70	0.0	0.0	30.0
	Slope	0.16	0.0	0.0	20.0
	Profile Curvature	0.12	0.0	0.0	15.6
	Northness	0.27	90.0	80.6	15.0
	PC1	0.69	0.0	0.2	13.7
	PC2	0.36	40.2	38.6	5.6

In terms of quantitative performance, the mean AUROCs (Figure 4.44) of PCA ModelA (0.77 in the nsCV and 0.68 in the sCV) and PCA ModelB (0.76 in the nsCV and 0.68 in the sCV) were quite similar, and approximately equal to the ST Model 1A (the best performing "physically plausible" model). Mean sCV AUROC being equal, PCA Model B had a very low IQR, thus resulting more spatially transferable than PCA Model A. Using the 42 rockfalls excluded from Configuration A models as an independent holdout test set, the AUROC was 0.81 for PCA ModelA.

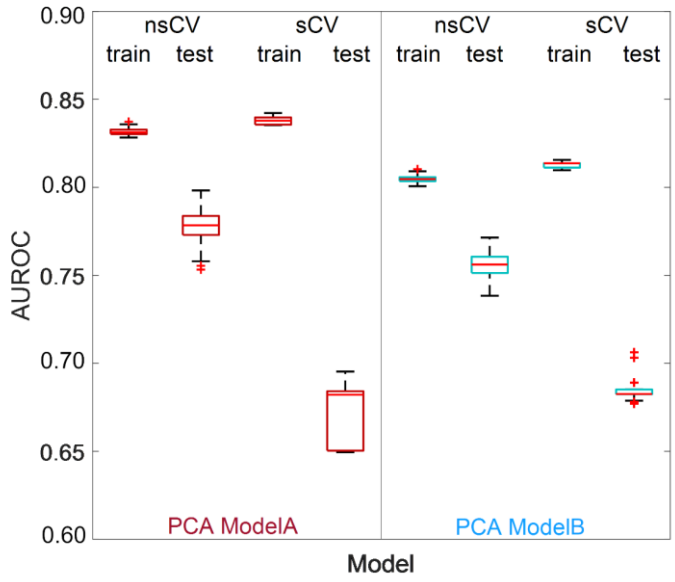


Figure 4.44 Boxplots showing the variability observed during the k-fold non-spatial and spatial cross- validation of the models PCA ModelA and PCA ModelB, regarding the area under the ROC curve (AUROC) values for training and testing landslide points.

4.7.2.6 Models comparison

In this section, a comparison (Figure 4.45 and Figure 4.46) of the spatial prediction patterns of six models - namely Blind-CLIMATE, ST Model 1A, GR-RedA, GR-RedB, PCA ModelA and PCA ModelB – will be provided.

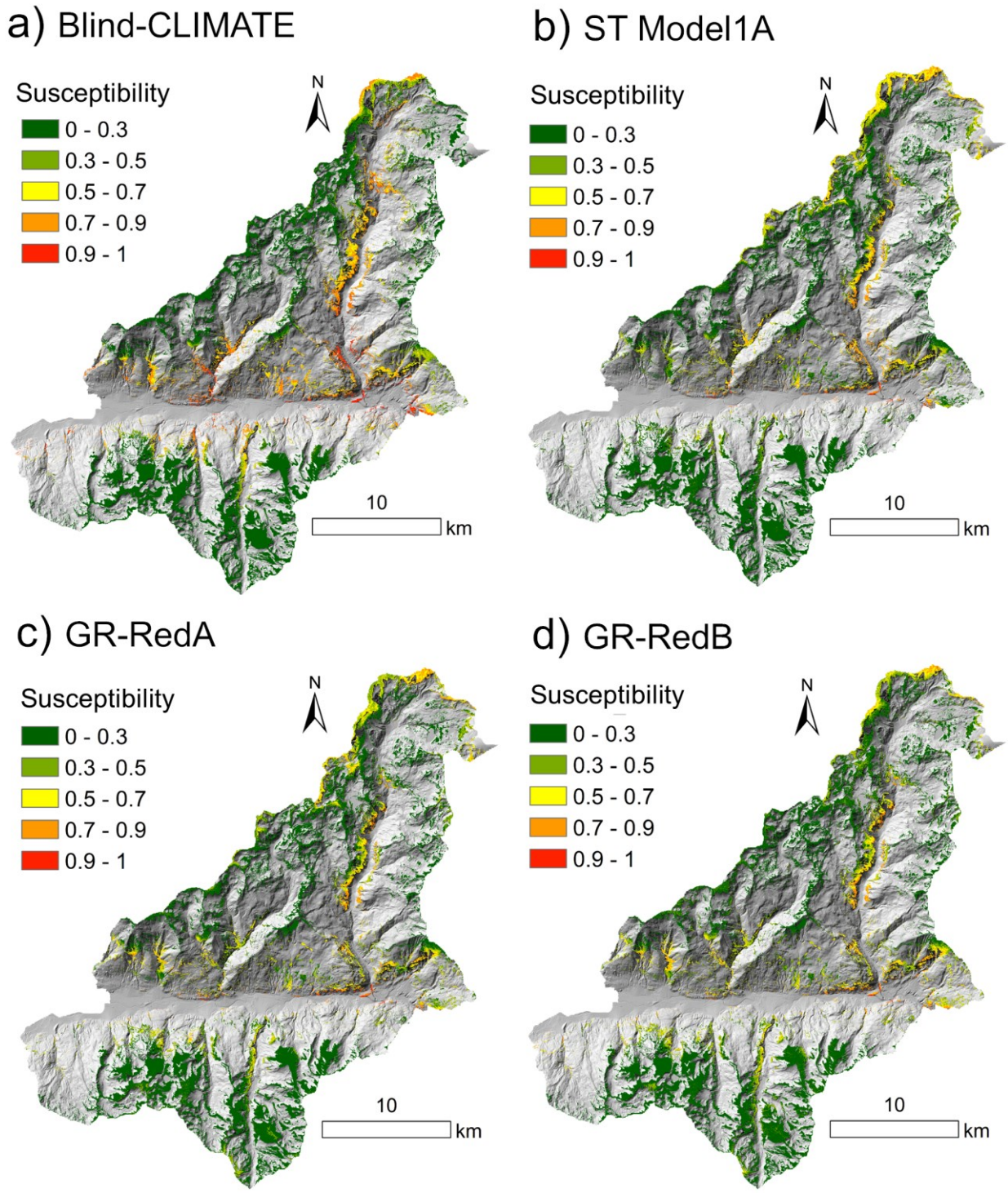


Figure 4.45 Reclassified rockfall susceptibility maps obtained from models: a) Blind-CLIMATE b) ST Model1A, c) GR-RedA and d) GR-RedB.

The spatial pattern of the Blind-CLIMATE model (Figure 4.45a) reflected the behaviour of elevation as the bias-describing predictor. Therefore, the highest susceptibility values were almost exclusively confined at the valley bottoms, close to roads and infrastructures, with a gradual decrease with elevation. This evidently indicated that the spatial pattern of susceptibility reflected the rockfall events collection procedures, rather than the environmental processes leading to rockfall occurrence. This problem was exacerbated by the addition to the model of the station-based climatic predictors EWIind and FT, which included elevation as regionalization covariate. With a total of three bias-describing predictors, it was therefore impossible to discern the data collection effects from the real susceptibility effects. The susceptibility spatial pattern of ST Model 1A (Figure 4.45b), GR-RedA (Figure 4.45c) and GR-Red B (Figure 4.45d) reflected the procedure adopted to reduce the inventory bias effects. The susceptibility gradual decreasing with elevation was not observed anymore, and it was interpreted that the high susceptibility values still visible at low medium elevation reflected the actual presence of active slopes.

Moreover, all the three maps indicated medium to high susceptibility values in correspondence of high peaks and rock walls (>2000 m a.s.l.) located in the northern and western part of the study area (head of the Valtournenche Valley), with a larger extent in the configuration A models (ST Model1A and GR-RedA) than in the GR-RedB model. In addition, the two models built with the grid-based climatic predictors (GR-RedA and GR-RedB) recorded a medium-high susceptibility values in some lateral

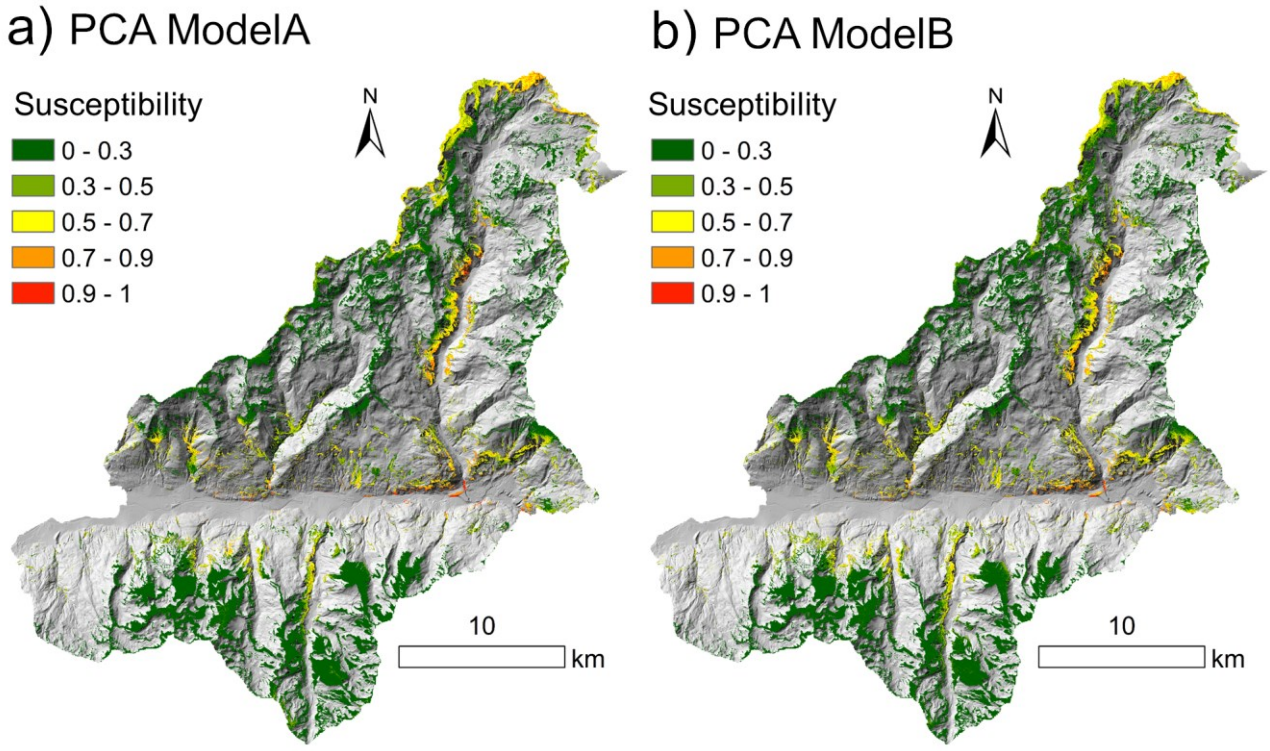


Figure 4.46 Reclassified rockfall susceptibility maps obtained from models: a) PCA ModelA b) PCA ModelB.

valleys located along the southern slope of the main E-W valley axis, not observed in the output map of the ST Model1A.

The PCA ModelA and PCA ModelB produced similar susceptibility maps (Figure 4.46a and Figure 4.46, respectively). They reflected the same patterns observed in the grid-based susceptibility maps (GR-RedA and GR-RedB). This is coherent with the fact that the principal components were calculated starting from the grid-based climatic predictors.

For all the models produced exploiting the visibility-related configurations it is necessary to highlight that the susceptibility values predicted outside the visibility mask fall in an area that was not covered (in case of Configuration A) or only marginally covered (in case of Configuration B) by sampling (training) points. Thus, the results over these areas may suffer from a more or less accentuated extrapolation degree, especially if those areas corresponded to predictors values beyond the extreme values covered by the predictors' CSF.

Table 4.19 reports the percentage of areas in each susceptibility class, for all the above-discussed models. For the models produced with the visibility configurations, the "high" and "very high" susceptibility classes were consistently reduced (about a half), if compared with the Blind-CLIMATE model. Moreover, the models including grid-based climate predictors were characterized by a further decrease of areas belonging to "high" and "very high" susceptibility classes respect to the model including the station-based predictors. Between configuration A and configuration B, the first presented more areas under the two highest classes than the second.

Table 4.19 Area percentages of each susceptibility class for the six compared models.

Susceptibility class	***Area %***					
	Blind-CLIMATE	ST-Model1A	GR-RedA	GR-RedB	PCA ModelA	PCA ModelB
Very Low 0-0.3	67.7	70.8	68.4	67.3	70.1	67.4
Low 0.3-0.5	8.9	12.7	14.9	16.3	13.7	16.1
Medium 0.5-0.7	9.5	10.3	11.0	11.0	11.0	11.8
High 0.7-0.9	10.6	5.2	4.9	4.8	4.5	4.2
Very High 0.9-1	3.4	1.1	0.8	0.5	0.8	0.5

The ten susceptibility difference maps prepared are summarized in Table 4.20 and discussed below.

VAR1-A. The susceptibility difference map between the ST-Model1A and the Blind-CLIMATE model (Figure 4.47a) showed the highest susceptibility variations, up to 0.9 (-0.9). As the bias-describing behaviour of altitude was smoothed through the visibility approach, negative variations were observed all along the valley bottoms, and positive variations were recorded for slopes located at high altitudes

and in remote secondary valleys. This susceptibility increase was very evident in the northern part of the study area, but present also in the southern part.

Table 4.20 Synthesis of the produced susceptibility difference maps and their associated names.

Difference Map Equation	*Map name*	*Difference Map Equation*	*Map name*
ST-Model1A — Blind-CLIMATE	*VAR1-A*	**GR-RedB — GR-RedA**	*VAR-AB*
ST-Model1A — TOPO-A	*VAR2-A*	**PCA ModelA — TOPO-A**	*PCA-VAR1-A*
GR-RedA — TOPO-A	*VAR3-A*	**PCA ModelA — GR-RedA**	*PCA-VAR2-A*
GR-RedA — ST-Model1A	*VAR4-A*	**PCA ModelB — TOPO-B**	*PCA-VAR1-B*
GR-RedB — TOPO-B	*VAR1-B*	**PCA ModelB — GR-RedB**	*PCA-VAR2-B*

VAR2-A. The susceptibility difference map between the ST-Model1A and the TOPO-A model (Figure 4.47b) showed susceptibility variations linked to the behaviour of the included station-based climatic predictors. In particular, positive values were recorded in the entire Valtournenche valley (i.e., the main N-S valley axis) as it was characterized by the highest values of WD-related threshold exceedance frequencies, which was the most important predictors in ST Model 1A. Positive values were also observed in some parts of the northern slopes facing the E-W main valley axis, which were characterized by the highest negative values of SWEmax, indicating consistent cumulated snow-melting.

VAR3-A. The susceptibility difference map between the GR-RedA model and the TOPO-A model (Figure 4.47c) showed a variation pattern very similar to map VAR2-A. For this map, positive values were also recorded along medium-high altitude slopes facing the main E-W valley axis.

VAR4-A. The susceptibility difference map between the GR-RedA model and the ST-Model1A (Figure 4.47d) reflected the difference in the spatial pattern between the station-based and grid-based climatic predictors. The utilization of the WD and EWIind grid-derived predictors led to a susceptibility increment for the GR-RedA model along both the northern and southern slopes of the E-W main valley. The slight negative values (up to a maximum of -0.3) observable at the north-western border of the study area could be imputed to the FT predictor, which had an increasing CSF for ST-Model1A but was penalized in the GR-RedA model.

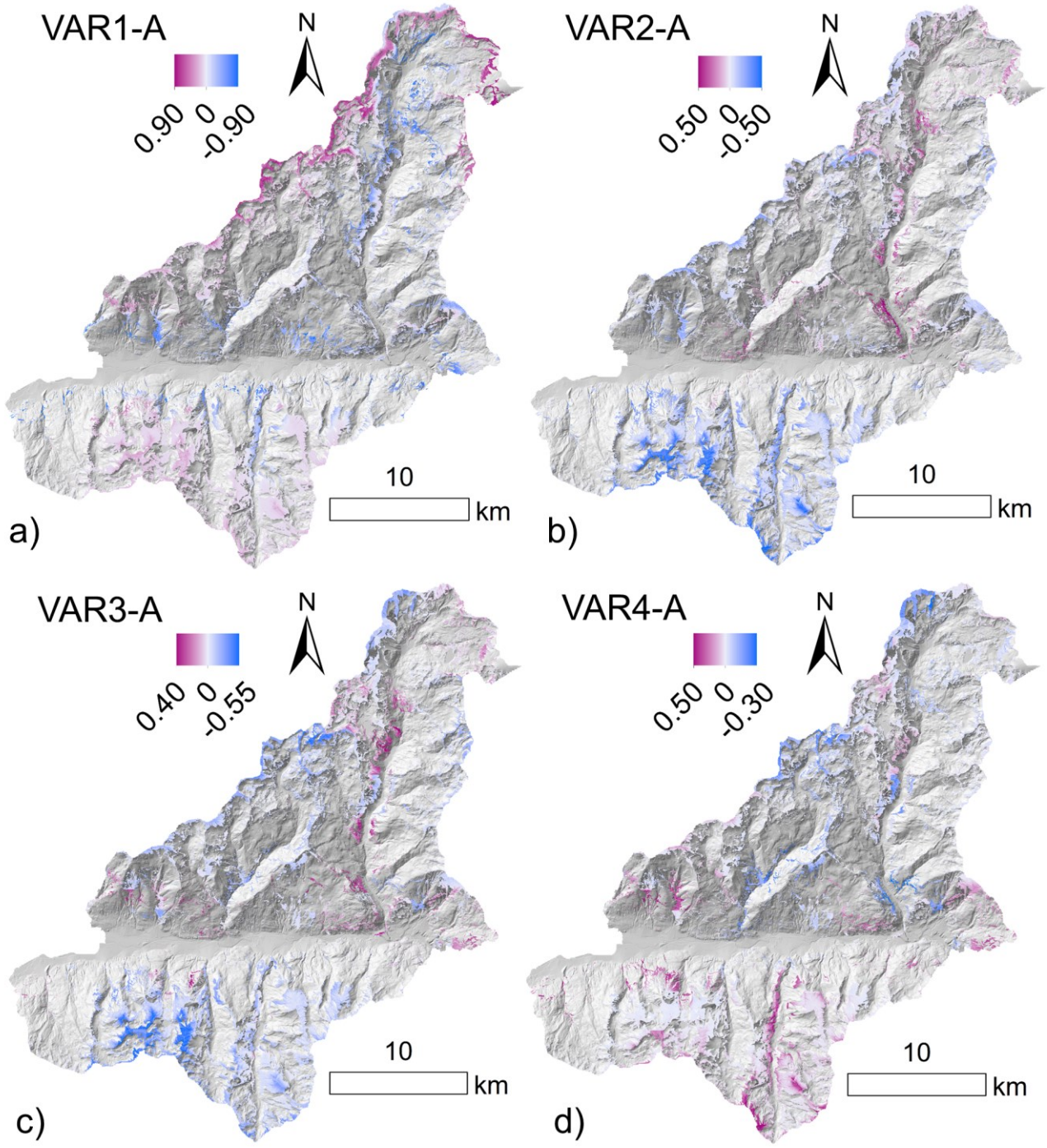


Figure 4.47 Difference maps for: a) VAR1-A=ST-Model1A — Blind-CLIMATE; b)VAR2-A=ST-Model1A — TOPO-A; c) VAR3-A=GR-RedA — TOPO-A and d) VAR4-A=GR-RedA — ST-Model1A.

VAR1-B. The susceptibility difference map between the GR-RedB model and the TOPO-A model (Figure 4.48a) nearly resembled both the spatial pattern and variation values range recorded in map VAR3-A.However, in comparison with the latter map VAR1-B shows more widespread positive values in correspondence of the north-eastern border of the study area.

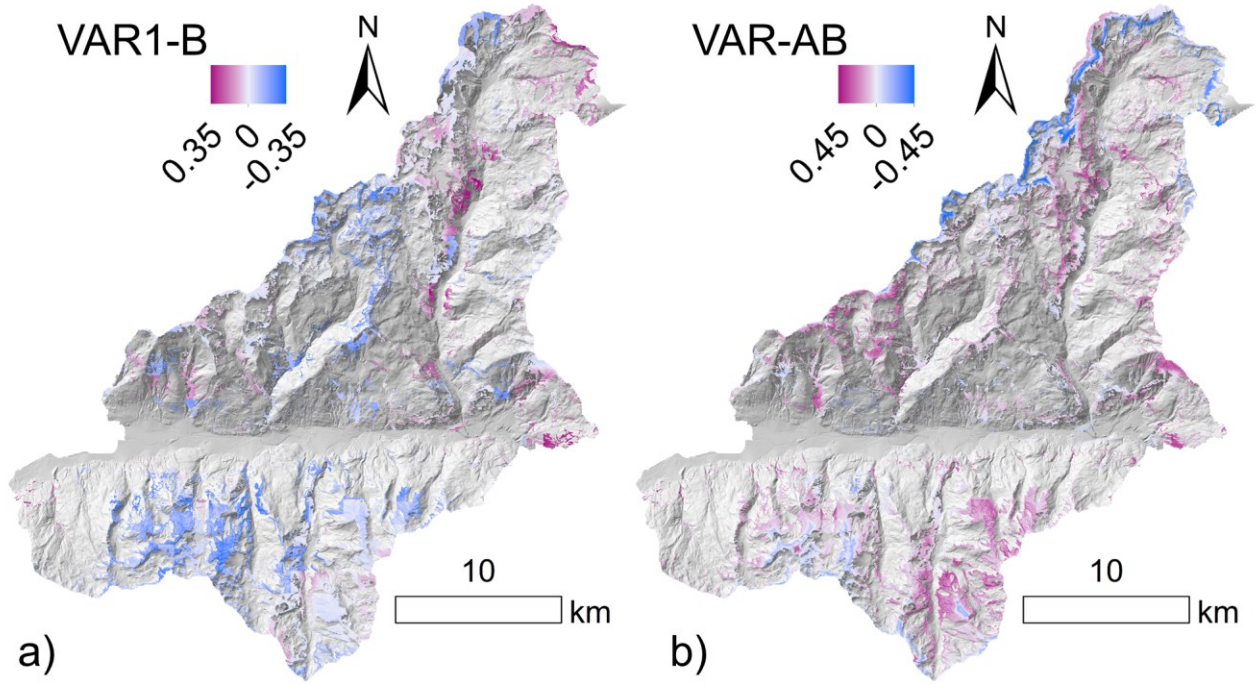


Figure 4.48 Difference maps for: a) VAR1-B= GR-RedB — TOPO-B and b) VAR-AB=GR-RedB — GR-RedA.

VAR-AB. The susceptibility difference map between the GR-RedB model and the GR-RedA model (Figure 4.48b) was characterized by widespread positive values, reasonably related to the steeper CSF characterizing the EWIind predictor in the GR-RedB model than in GR-RedA model. In the GR-RedA model, WD and SWEmax predictors were characterized by steep increasing CSFs, while, in GR-RedB model, WD was penalized and SWEmax had a very weakly inclined function. This different behaviour explains the negative values along the north-western and north-eastern borders of the study area, where these two predictors actually showed their maximum values.

PCA-VAR1-A and **PCA-VAR1-B.** The susceptibility difference maps between the PCA related models and their TOPO model counterparts were very similar (Figure 4.49a and Figure 4.49c) and reflected the distribution of the highest threshold exceedance frequencies for EWIind, WD and FT and of the cumulated snow melting processes. Overall, the inclusion of climatic predictors in terms of principal components had led to a susceptibility increment in the Valtournenche Valley and on the northern slope facing the E-W main valley axis, while a susceptibility decrement could be observed especially at high altitudes of the southern slope facing the E-W main valley axis.

PCA-VAR2-A and **PCA-VAR2-B**. The susceptibility difference maps spatial patterns between the PCA related models and GR-RedA and GR-RedB models (Figure 4.49b and Figure 4.49d, respectively) could be basically explained by the absence of the SWEep predictor for model GR-RedA and of the SWEep and FT predictors for model GR-RedB.

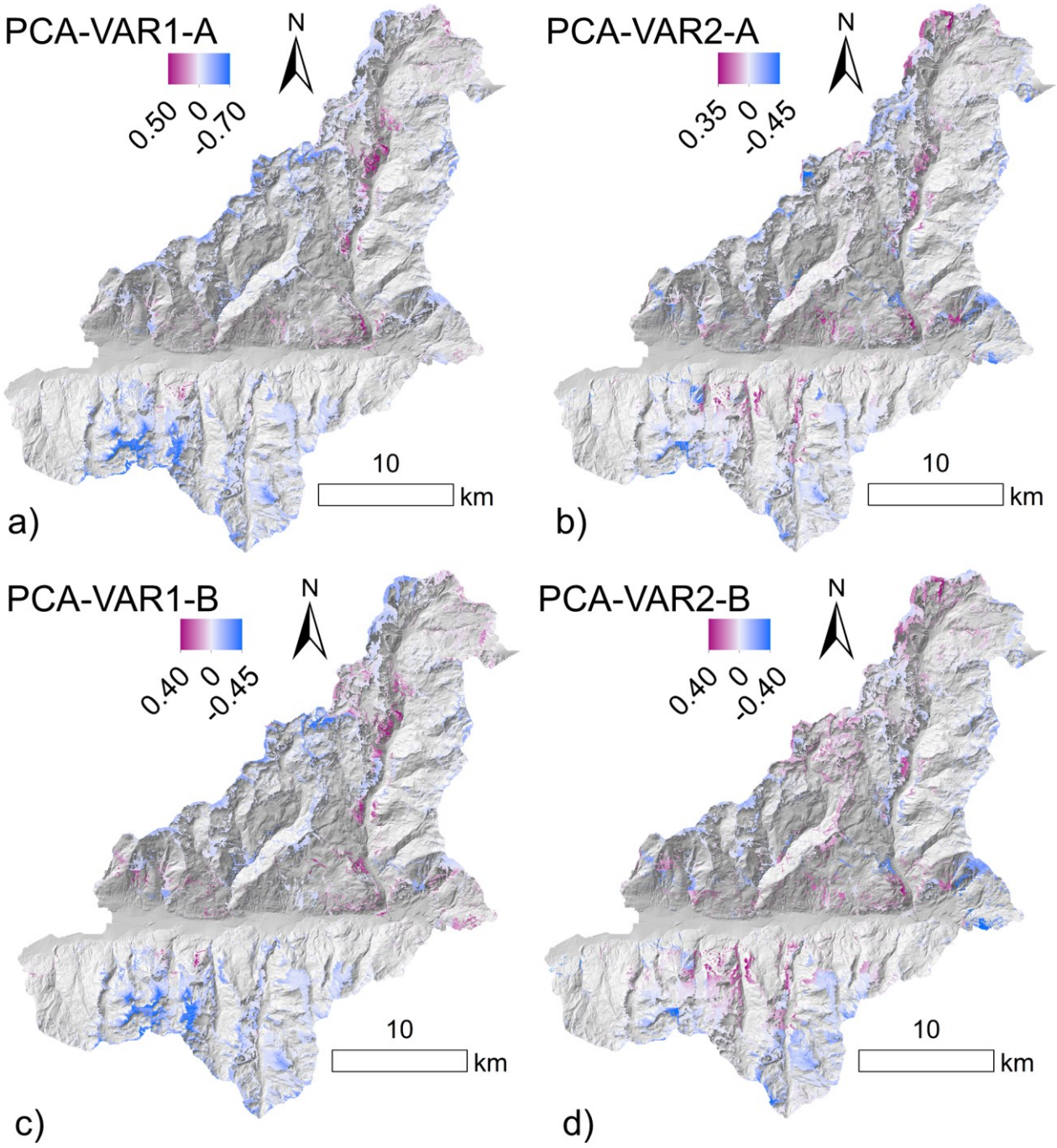


Figure 4.49 Difference maps for a) PCA-VAR1-A= PCA ModelA — TOPO-A; b) PCA-VAR2-A= PCA ModelA — GR-RedA; c) PCA-VAR1-B= PCA ModelB — TOPO-B and d) PCA-VAR2-B= PCA ModelB — GR-RedB.

4.8. Discussion and future perspectives

The first phase of the work explored four climatic indices as possible triggering and preparatory factors for rockfall instabilities in a sub-area of the Aosta Valley Region, both considering short-term and long-term durations. Gunzburger et al. (2005) outlined that the distinction between triggering and preparatory factors, defined in terms of their action time frame, should not be intended as an absolute dichotomy, as they act in a continuous transition. This concept seems to be confirmed by the present work, as several rockfalls resulted to be associated to more than one duration corresponding to not-ordinary conditions, and to a combination of different climatic indices acting at different temporal scales. In such cases, it is almost impossible to state which condition (long-term index or short-term index) was responsible for the final collapse. However, the definition of not-ordinary conditions, the identification of critical combinations of indices and the recognition of thresholds can lead to a better understanding of the mechanisms driving such events, helping in defining the most critical situations.

The historical rockfall inventory of Aosta Valley Region was crucial for the whole analysis. When possible, not neglecting to record date and time of each rockfall event is essential for the calculation of specific climatic indices and thresholds. However, although rich and suitable for the first phase of the study, the analysis made on the location of source areas suggested a negative bias in event record acquisitions at high, poorly frequented elevations. This is a common and widespread problem, which leads to an incomplete (to an unknown extent) inventory (e.g., Petschko et al., 2014).

The highest rockfall frequency occurred in spring (March to May) and a secondary one in winter, with its main peak in January. Macciotta et al. (2015) found a comparable peak in January, explained by freeze-thaw cycles occurrence. Frayssines and Hantz (2006) found a similar annual cycle for rockfall occurrence in the area surrounding Grenoble (France), but with the primary peak in winter rather than spring. Compared to this study, the authors focused on larger rockfall volumes (mainly > 50 m3) and lower altitudes (from 200 to 2000 m a.s.l.), thus they investigated events characterized by a lower occurrence of freeze-thaw cycles in spring.

The comparison between climatic conditions recorded before each rockfall occurrence and ordinary (i.e., typical) climatic conditions for the area followed the approach of Paranunzio et al. (2015). Differently from these authors, in this study the cumulated precipitation index was designed to include the water inputs coming from snow melting processes and it was renamed as effective water input index. In addition, other processed indices were explored, namely wet and dry episodes and freeze-thaw cycles. Conversely to D'Amato et al. (2016), who found that short term rainfall gives the higher rockfall frequency, the occurrence of not-ordinary conditions for the short-term rainfall (STR) index showed

poor correlation with rockfall occurrence. This difference may be explained by the more limited rockfall volume range studied by D'Amato et al. (2006), which is mainly below 1 m^3. In the Canadian Cordillera, Macciotta et al. (2017) suggested that snowmelt can explain a peak in rockfall record in February-March. Accounting for snow melting inputs was possible by developing a simple method based on degree-day-factor and snow data conversion. The obtained representative degree-day-factor of 3.5 mm C−1 day−1 resulted to be coherent with typical literature values, which range from 1.6 to 6.0 mm C−1 day−1 (USDA-NEH, 2004; He et al., 2014). The adopted correction method introduced some simplifications regarding the complex dynamics of snow related processes. However, the results obtained demonstrated the primary importance and utility of snow derived water inputs when investigating reasons for rock masses instabilities. Therefore, in agreement with findings of several authors (e.g. Nishii et al., 2013; Crosta et al., 2014), the results of this study strengthen the need for in-depth research on the snow dynamics role on rockfall occurrence.

For threshold construction, the lowest critical percentile value was used when not-ordinary conditions were exceeded for more than one duration. This choice represents a conservative approach, which is reasonably preferable when dealing with natural hazards. In addition, the lowest critical value could correspond to residual geomechanical conditions and therefore it could consider the possibility of rockfall occurrence as secondary reactivations (i.e., linked to less intense triggers/preparatory factors). The lower dispersion obtained with normalized thresholds in comparison to those derived from the simple Intensity-Duration formulation confirms the utility of this approach in areas characterized by the juxtaposition of different climate types. This result agrees with findings of numerous previous studies (Guzzetti et al., 2007, 2008; Peruccacci et al., 2017; Leonarduzzi & Molnar, 2020).

Despite the adoption of a normalization approach, the FT threshold construction was difficult due to high dispersion of the input data. Indeed, not ordinary conditions in term of freeze-thaw cycles can be very different depending on altitude, also considering that T is a continuous parameter and rockfalls can be the expression of the loss of the equilibrium between the rock mass system and the external thermal conditions. In this study, the temperature variability with altitude was implemented by adopting an annual ATLR of −0.53°C/100 m, which is coherent with the annual lapse rates ranging from −0.54 to −0.58°C/100 m typical for various mountain regions in France and Italy (Rolland, 2003). As shown by Nigrelli et al. (2018) and Kirchner et al. (2013), the application of a monthly or sub-daily ATLR could guarantee a detailed representation of temperature variation dynamics (temperature inversion phenomena, aspect and slope influence, vegetation and snow cover indirect effect through incoming solar radiation) and possibly an improved modelling of FT cycles. However, it should be considered that

the final aim of this study was not to analyse a particular rockfall or meteorological event, but to produce an overall analysis of the influence of different climate indices on rockfall occurrence and to develop thresholds extendable on a sub-regional scale. Thus, the annual (generally accepted) ATLR approach on temperature variations was considered a reasonable compromise at this investigation level. Also, few attempts to translate freeze-thaw cycles in quantitative terms and thresholds are present in the literature (D'Amato et al., 2016) and the results can be considered promising and with scope of improvement.

During Spring, a high number of events was related to water inputs (both in the form of precipitation and WD episodes) plus FT cycles not ordinary conditions. Therefore, these occurrences could be explained through the processes described by Matsuoka (2008) and Draebing and Krautblatter (2019). According to these authors, rockfalls can occur when water is available in cracks and joints (guaranteed by rainfall not ordinary conditions) and temperature conditions allow for frost wedging or ice segregation (guaranteed by FT cycles not ordinary conditions). Also, FT cycles alone are recognized as a highly frequent preparatory-triggering factor. Despite the amplitude of thermal fluctuations of each cycle was not investigated in this work, in terms of processes, these occurrences could be related to repeated rock thermal stresses as described by Matsuoka (2019).

Only six rockfalls could not be associated to not ordinary conditions for the investigated climate indices. However, a preliminary exploratory analysis of these rockfalls suggested that they could be potentially referred to the occurrence of high temperatures and wide daily temperature gradients. Based on the elevations at which they took place, only a single rockfall could be referred to permafrost degradation (Matterhorn, 3735 m a.s.l., on 04/08/2003) as described by several authors working on high elevation rockfalls (Allen and Huggel, 2013; Bodin et al, 2015; Paranunzio et al., 2016; Nigrelli et al., 2018). The event occurred at a medium-high elevation (1850 m a.s.l.) on 05/06/2009 could be possibly related to local episodes of late spring snow melting, as suggested by Paranunzio et al. (2016). The event recorded on 09/07/2017, at an elevation of 1175 m a.s.l., showed strong anomalies both in terms of high temperatures and temperature gradients. Due to its elevation and date of occurrence, it fits well with the mechanism of thermal weathering described by Collins and Stock (2016) and Collins et al. (2018), as well as the thermal related sub-critical cracking described by Ishikawa et al. (2004). The same could be assumed for other three events, even if they had a weaker signal in terms of the effect of high temperatures and gradients. These few events are evidence of the growing effect of thermal fluctuations and extremely high temperatures on rock slope instabilities in the alpine context. Having in mind the projected global warming, these preliminary results are therefore an indication that specific studies to

enhance the knowledge related to these processes is necessary not only in arid and semiarid environments, but over Alpine regions too.

The second phase of the work was aimed at translating the not-ordinary climatic processes and the correspondent empirical thresholds in synthetic climatic variables to be used as predictors in a statistically-based rockfall susceptibility model. Different authors addressed the inclusion of the climatic component in their susceptibility models, adopting various approaches: (i) by treating precipitation in the form of mean annual rainfall, mean monthly rainfall and rainy days frequency (e.g., Broeckx et al., 2018; Chen and Li, 2020; Fang et al., 2020; Nahayo et al., 2019; Nhu et al., 2020); (ii) by developing event-based landslide susceptibility assessments, recognizing cumulated rainfall in a defined period of time as an important predictor of the model (e.g., Kim et al., 2015; Gassner et al., 2015; Knevels et al., 2020); (iii) by coupling in different ways (e.g., contingency matrices, models multiplication) susceptibility models exclusively based on the traditionally morphometric predictors with rainfall thresholds exceedance (Segoni et al., 2015, 2018) or with antecedent cumulated rainfall and soil saturation degree (Bordoni et al., 2020) in an early warning perspective; (iv) by introducing a certain number of synthetic variables expressing the return period of different rainstorms with specific given total rainfall amount and durations (e.g., Catani et al., 2013). However, all the above-mentioned works were focused on shallow landslides, while rockfalls are often disregarded in terms of sources individuation, on behalf of runout analyses (Michoud et al., 2012). In this study, as a critical duration or amount for the investigated climatic processes (i.e., effective water inputs, wet and dry cycles and freeze-thaw cycles) is neither a priori known for rockfalls, nor given by the Regional Civil Protection Office, a similar rationale to Diskshit et al. (2020) and Camera et al. (2021) was considered the most suitable and therefore adopted. It consisted in calculating the mean annual threshold exceedance frequency for each one of the recognized climatic processes influencing rockfall occurrence and using these metrics as predictors for the susceptibility modelling phase. The climatic predictors creation was based both on station-based and grid-based datasets. It can be stated that the grid-based procedure, despite based on an interpolation step at its origin (refer to Section 4.3.3.2), may be deemed more reliable than the station-based approach as each pixel has its own hourly time series and consequently its own independently calculated threshold exceedance frequency (TEF_a,) for each specific index. In the case of the station derived predictors, each pixel has an inferred TEF_a, coming from a regionalization which, although with good performance results, may miss the local variability and microclimate differentiation, recording exclusively the strongest signal (e.g., FT correlation with altitude). Nevertheless, hourly grid-based meteorological datasets are rarely available; the applicability of the station-based procedure in other alpine contexts is much more achievable and remains a valuable option, especially if time and resources

could be spent in finding the most suitable regionalization technique and performance metrics would return satisfying results. Moreover, it has to be considered that the code for the elaboration and calculation of TEF_a for the grid-based dataset had taken some days to run, conversely the calculation of TEF_a with station-based dataset was immediate, although some time was required to test the different regionalization techniques. Finally, to maximize the regionalization performance for station-based TEF_a, some morphometric covariates needed to be included in the process: this lead to multi-concurvity problems during the susceptibility modelling phase.

The main advantages of the threshold exceedance frequency approach (Camera et al., 2021) are its conciseness, versatility and non-stationarity (i.e., updatability). It is concise as only one variable for each process is used, instead of several variables describing different specific rainfall amounts, durations or return periods. It is versatile because potentially derivable from all types of meteorological datasets, depending on their availability in the study area of interest (i.e., from meteorological stations, grid datasets, radar and satellite products). It is updatable since they may be calculated for different reference temporal periods and for future climatic scenarios. Moreover, only the 0.3% of the landslide susceptibility related literature dealt with climate processes, and the majority of them focused only on rainfall (Reichenbach et al., 2018). Temperature variations, also in the form of freeze-thaw cycles, were never included in a rockfall susceptibility assessment. Messenzehl et al. (2017), who modelled rockfall susceptibility with Random Forest in the Swiss Alps, identified the regional permafrost distribution as the most important variable controlling the spatial rockfall activity and imputed the observed rockfall clustering at a specific altitudinal window (> 2500 m a.s.l.) to a highly effective action of freeze-thaw cycles. However, they did not include freeze-thaw cycles directly as a variable in the model. Conversely, even if wet and dry cycles are undoubtedly recognized as a concerning weakening process in rock masses, they are actually only investigated at the laboratory level (e.g., Van der Hoven et al. 2003; Torres-Suarez et al. 2014; Zhou et al. 2017; Yang et al., 2018, 2019). Therefore, the innovative imprint of this study relies both on the subject under analysis (i.e., rockfalls) and on the exploration of several climatic processes consistent with rock slope instabilities.

The majority of the landslide susceptibility studies available in literature assess the susceptibility model quality and meaningfulness exclusively on the model quantitative performance. Any problem concerning the positional accuracy and spatial representativeness of the inventory is often disregarded (Reichenbach et al., 2018; Steger et al., 2021). Ignoring these inventory-related issues usually lead to extremely good performances but altered modelled associations and maps (Steger et al., 2017). Steger et al. (2021) suggested to discern susceptibility effects and data collection effects, as the distribution of

inventoried landslides depends on the methodological approach adopted for data collection. Although Aosta Valley had implemented an efficient procedure to integrate landslide public reports, remote sensing data and Forest Corps constant monitoring (see Section 4.3.1), an inherent bias towards an overrepresentation of damage-related events and an underrepresentation of remote areas is almost unavoidable. This is common in inventories obtained from public administrations (Guzzetti et al., 1999). The reason is that the inventories are compiled for operational or Civil Protection purposes, and not for capturing all the possible environmental combinations leading to instability. In the case of rockfalls, this bias may be exacerbated as the occurrence of this type of instabilities in remote sub-vertical inaccessible rock walls is common. To investigate these frequent shortcomings a "Blind" modelling approach was initially carried out, ignoring on purpose the underlying inventory bias. Even if this modelling procedure suggested an excellent rockfall discrimination capability in terms of performance ($0.8 \leq AUROC \leq 0.9$), the models showed an implausible and distorted representation of rockfall susceptibility in terms of both predictors' behaviour (e.g., elevation, EWIind, FT and SWEep) and susceptibility spatial pattern. In detail, elevation related to susceptibility with a decreasing linear CSF and the most susceptible areas were relegated to the main valley bottoms, while high altitude areas far from roads and infrastructure were characterized by very low predicted susceptibility. These results were coherent with other works for which a similar modelling framework to deal with possible inventory bias was adopted (e.g., Steger et al., 2017, 2021; Bajni et al., 2021b). These findings also make questionable the frequent use of distance from roads and distance from river predictors in landslide susceptibility modelling (Reichenbach et al., 2018), as their role could exclusively consist in a bias-reinforcement.

Researchers are recently focusing on finding strategies to limit these extremely common inventory bias issue. A possible strategy consists in the adoption of Generalized Linear Mixed Models (GLMM), in which the bias-describing predictors are specified as a random intercept, in order to isolate the model variations related to the bias from the fixed effects of the other variables (Steger et al., 2016b, 2017). Another approach is based on the modification of the sampling strategy for non-landslide points, limiting the modelling area to the one monitored by administrations and technicians responsible for inventories compilation and update (Knevels et al., 2020; Bornaetxea et al., 2018). Another possibility is to exploit the inventory bias to shift the modelling subject from landslide susceptibility to areas likely affected by damaging landslides and consequent interventions (Steger et al., 2021). In this work, a sampling strategy based on the creation of "visibility" masks (i.e., recognition of reduced eligible areas), similar to Knevels et al. (2020) and Bornaetxea et al. (2018), was preferred for several reasons. Firstly, Steger et al. (2021) demonstrated that in presence of numerous variables with the double role of real rockfall susceptibility-related predictor and bias reinforcing predictor (i.e., rockfalls overrepresentation towards roads and

infrastructures), the GLMM failed in predicting susceptibility in a reliable way. In this work, these predictors were at least three: namely, elevation, EWIind (especially the station-based one) and FT. Secondly, the approach proposed by Steger et al. (2021) consisting in the change of the subject to be modelled into a landslide intervention index, poorly matches with the aim of the present study, dealing with the inclusion of non-stationary predictors in the susceptibility model. Indeed, considering the actual spatial pattern of damaging landslides, and not the underlying physical processes, represents a support for the present-time risk managing activities, but it is not helpful for the mid- and long-term planning of land uses. Moreover, the same rationale of excluding quaternary deposits from the modellable area when dealing with rockfalls (and *vice versa* excluding rocky outcrops when dealing with shallow landslides) should be adopted. Indeed, it may be argued that, if the interest is just on areas potentially damaged by landslides, the study area should be reduced around infrastructures, where actually damages can occur. The visibility approach resulted to be valuable for smoothing the issues related to the inventory bias. However, relationships derived and validated in the visibility domain are assumed to be valid in completely unseen areas, where they could result in extrapolations. The two "visibility" configurations had comparable AUROCs and similar coverage of the area characteristics, in terms of geo-environmental predictors combinations. The latter can be evaluated with both qualitative observations and quantitative metrics, e.g., by Multivariate Environmental Similarity Surface (MESS, Elith et al., 2010). Configuration A was deemed more rigorous and offered the possibility of retaining a hold-out test (17% of the total sample). Configuration B aimed at maximising the possible geo-environmental combinations representable in the model; the idea was to couple an equal number of non-rockfalls to rockfalls occurred outside the visibility mask, thus matching the "randomness" or "accidentalness" of outside-rockfalls reporting. However, outside the "visibility" mask, rockfall phenomena were underrepresented in the inventory, thus it was not guaranteed that the selected correspondent non-rockfalls points occurred in non-active areas. Despite the different strategies adopted and discussed above, researchers should make an effort to open a communication channel with local administrations to get a synergic change in the landslides collection approach underway. It is of vital importance that landslide inventories are able to reflect the actual distribution of instabilities not only for a mere scientific interest, but to provide benefits to the mountain communities through an informed future environmental planning. This becomes even more crucial if the analyses include non-stationary predictors (e.g., climate related predictors and land use).

Another concerning drawback in landslide susceptibility modelling is neglecting the modelled predictors' behaviour physical plausibility, which is particularly crucial when dealing with process-driven predictors (Camera et al., 2021; Bajni et al., 2021b). Including predictors that improve the

quantitative performance of the model but express an implausible physical and geomorphological behaviour may lead to misleading susceptibility maps and process interpretation. This may be exacerbated when dealing with sophisticated Machine Learning models, which usually do not offer an easy interpretation of predictors behaviour and thus are considered successful only based on model quantitative performance (Goetz et al., 2011; Brenning et al., 2012b). For these reasons the component smoothing functions (CSF) of both morphometric and climatic processes were carefully analysed for all the models produced. The morphometric predictors remained quite unvaried through the different models and visibility configurations tested. The climate-related predictors CSF slightly varied moving from the visibility configuration A to the visibility configuration B. Also, they varied when alternatively using the station-derived and grid-derived climate variables. The not-physically plausible climatic predictors were excluded from the models, even if their presence increased the model quantitative performance. Some common characteristics could be however inferred. For example, EWIind and WD were alternatively selected as the most important, not-penalized, physically plausible predictors (with mDD% near 10% each), followed by SWEmax (mDD% ranging from 3% to 5%). In some models, the EWIind CSF was characterized by a near sigmoidal shape in its initial part, as already observed in other works including cumulated rainfall-related variables in landslide susceptibility models (e.g., Tanyas et al., 2021). The predictor SWEep was always represented by a physically distorted CSF, while FT was basically always penalized as its behaviour was masked by elevation. In presence of concurvity (as between FT and elevation), the non-selection of penalized terms may imply that other covariates have similar and stronger mathematical relationships with the response variable, and not that the penalized terms are not related to the investigated phenomena (Laceby et al., 2016). It is interesting to note that elevation, in both Configuration A and B, presented a highly non-linear CSF, characterized by a raising tail at the highest elevations (>2500 m a.s.l.), also contributing to increase susceptibility in correspondence of the high mountain peaks. This is particularly evident in the northern part of the study area. This behaviour seemed to be coherent with the findings of Messenzehl et al. (2017), who found an intense rockfall activity linked to permafrost probability at similar altitudes (>2900 m a.s.l.) and imputing it to permafrost degradation and consequent water seepage and hydrostatic pressure variations. EWIper and EWIind predictors expressed threshold exceedances due to intense precipitation, but in two alternative ways, therefore showing different spatial patterns. As a result, their role within the model was different. EWIper always showed a physical implausible behaviour, thus the models including EWIind were preferred. However, EWIind showed a strong correlation with altitude, the most bias-reinforcing morphometric predictor. As it was not possible to quantify numerically the bias removal

extent, doubts remain on the absolute better efficiency of EWIind in comparison to EWIper. This rationale particularly applies to the station-derived predictors.

Finally, it is a common practice to carry out a multicollinearity diagnosis of predictors before using them as independent variables in a susceptibility model, both when dealing with logistic regression and complex Machine Learning models (Chen W. et al., 2017; Nohani et al., 2019; Roy et al., 2019; Yi et al., 2020; Arabameri et al., 2020; Alqadhi et al., 2021). However, a standard procedure to select which predictors should be excluded, or retained, in case of high collinearity is missing. Segoni et al. (2020) developed a Random Forest landslide susceptibility model including contemporaneously the geological predictor with six different classifications. The authors pointed out that advanced Machine Learning algorithms are able to handle predictors collinearity and, also, it is widely accepted to simultaneously use morphological and hydrological strongly correlated variables. In the case of GAMs, introducing in the model predictors characterized by high concurvity values may cause, as for collinearity, unstable estimates and predictors behaviour. However, the adoption of the penalization-based (shrinkage) variable selection allows to set to zero those terms that are involved in multi-concurvity issues (Laceby et al., 2016; Bagalwa et al., 2021). To test a possible alternative strategy both to reduce concurvity and to increase model parsimony, a Principal Components Analysis (PCA) was carried out. The grid-derived climatic predictors were used as the PCA inputs and subsequently replaced by the derived principal components as susceptibility model predictors. Messenzehl et al. (2017; 2018) adopted a similar approach, carrying out a PCA to model rockfall activity in the Swiss Alps. The two PCA-based models showed encouraging results, as the principal components CSF had an explainable physical meaning. The principal components showed a quite high importance in the model and had the advantage to consistently decrease concurvity below the 0.8 threshold. On the other hand, it should be also highlighted that climatic predictors are inherently non-stationary, while elevation is a static variable, thus concurvity between them could vary in the future, as threshold exceedance patterns evolve. Moreover, PCA-related models embody the quality of parsimony, but they may be less immediate to interpret as it is inherently more difficult to discern the extent of the influence of each single climate process in defining the component behaviour and importance.

Model performance was tested both through a non-spatial and a spatial k-fold cross validation. In the visibility approach related models, the AUROC derived from the sCV showed a consistent drop of about 0.10 if compared with the nsCV counterpart, whose results could be considered indices of good performance. Moreover, the interquartile range of the AUROC estimates coming from the different sCV runs was quite high (except for the PCA ModelB), meaning a low spatial transferability of the modelled

relationships between climatic predictors and susceptibility. Although a spatial CV is usually deemed as the best choice when dealing with spatial data (Brenning et al., 2012b), it may have some issues when dealing with climatic data, especially in presence of mountain related microclimatic variations, which is the case of the study area examined in this work. Spatial partitioning may induce extrapolations by restricting the ranges or combinations of predictor variables available for model training, thus possibly leading to large prediction errors (Roberts et al., 2017), not accountable on the algorithm capabilities but on the training points selection strategy. Indeed, the threshold exceedance frequency varies according to the location and its microclimate; removing continuous portions of the landscape to use the related rockfalls as test set may cause the exclusion of a range of values from the construction of the climatic predictor CSF. An independent holdout test set, in a neighbouring region with a similar altitudinal and geographical extent, could be a better option to test the spatial transferability of the models including the climatic component.

The inclusion of dynamic predictors in landslide susceptibility modelling, making it adaptable to map the future evolution of instability patterns is of primary importance (Reichenbach et al., 2018; Camera et al., 2021). Climate change projections demonstrate that air and ground temperatures will increase through the end of the 21st century (Gobiet et al., 2014; Bender et al., 2020). In addition, an elevation-dependent enhanced warming (almost double) has been observed in mountain regions (IPCC, 2007; Auer et al., 2007; Pepin et al., 2015). As a possible effect, cryogenic processes and associated triggered rockfall would be shifted to higher elevations (Draebing, 2021). Moreover, according to IPCC (2013), the frequency of heat waves in large parts of Europe (including the Alps) has increased since the middle of the 20th century and are expected to increase in future (Gobiet et al., 2014). Also, global warming is likely to enhance the frequency and magnitude of extreme precipitation events (Rajczak et al., 2013; Gobiet et al., 2014; Stoffel et al., 2014), although future precipitation projections and models are characterized by substantial uncertainties (Fischer et al., 2013). Overall, future precipitation patterns indicate less precipitation in summer, up to -20%, particularly south of the Alps, and enhanced precipitation in winter, up to +10% by the end of the 21st century (Gobiet et al., 2014). If these changes would take place, it will also affect the occurrence of wet and dry cycles. Regarding snow dynamics, projected changes include variations in the rain-to-snow fraction, a seasonal shift to an earlier snowmelt initiation date and a general decrease in the length of the snow cover season (Beniston et al., 2018; Citrini et al., 2020; Javadinejad et al., 2020; Bender et al., 2020; Vorkauf, 2021; Chiarle et al., 2021).

To actually make the susceptibility maps non-stationary and including the above mentioned variations and extremization of climate processes, two possible and complementary future developments of the

work are addressable: (i) producing future susceptibility scenarios including the climate change component and ideally testing different emission scenarios or RCPs (IPCC,2014) for long-term environmental and infrastructure planning and (ii) implementing the statistically based susceptibility models in an early warning system for Civil Protection purposes. For the first implementation all the not-penalized and physically plausible climatic predictors (EWIind, WD, FT, SWEmax) should be defined for future reference periods, and the susceptibility maps updated on the basis of the previously defined smoothing functions. For the second implementation, the mean annual exceedance frequencies need to be transferred in an operative perspective, to include both short/mid-term precipitation and temperature forecasts (or archived analysis products for model development and training). In order to include antecedent conditions and the preparatory role of the climatic predictors (i.e., the long durations terms of the defined thresholds), observations or reanalysis data could be exploited if available for the time periods of interest.

Chapter 5

Conclusions

A systematic procedure capable to examine and enhance the role of geomechanical and climatic processes in rockfall susceptibility modelling was established. Two case studies were considered in the Italian Alps: Valchiavenna, in the Lombardy Region and the Mountain Communities of Mont Cervin and Mont Emilius, in the Aosta Valley Region.

Novel spatially distributed predictors specifically suited for rockfall occurrence in Alpine environments were developed. For the Valchiavenna case study, outcrop-scale geomechanical properties namely Joint Volumetric Count (Jv), rock-mass weathering index (Wi) and rock-mass equivalent permeability (Keq) were produced. Starting from an available dataset of rock mass geomechanical data (e.g., joint spacing, aperture, roughness, Schmidt Hammer rebounds), 25 new geomechanical survey locations were optimally selected thanks to the application of the Spatial Simulated Annealing sampling method. The field collection of new geomechanical data in these locations allowed to reduce kriging variance associated to Jv and improve the representativeness of the geomechanical dataset in terms of Multivariate Environmental Similarity Surface (MESS). Subsequently, to produce a map of each property as reliable as possible, both in terms of performance and representativeness of the geomechanical setting of the area, several geostatistical, deterministic and regressive regionalization methods were tested. Geomechanical predictor maps were obtained applying an ordinary kriging for Jv and Wi (NRMSE equal to 13.7% and 14.5%, respectively) and by means of Thin Plate Splines for Keq (NRMSE= 18.5%). For the Aosta Valley case study an approach to derive and include climate variables, reproducible exploiting both station-based and grid-based meteorological dataset, was developed. Firstly, robust cause-effect relationships between climate processes and rockfall eventswere identified and then translated in critical empirical thresholds for rockfall occurrence. In detail, three climate indices representative of the processes were identified: namely the cumulated amount of effective water inputs (EWI, including both rainfall and snow melting intake), the cumulated number of wet-dry cycles (WD) and the cumulated number of freeze-thaw cycles (FT). The 95% of the rockfalls occurred in severe (or not ordinary) conditions for at least one among the EWI, WD and FT indices. Subsequently, the mean annual threshold exceedance frequency for each climate index was calculated and adopted as a predictor for the susceptibility analyses. Threshold exceedance frequency patterns deriving from the station-based

and grid-based datasets showed similar spatial patterns; nevertheless, some differences in maximum and minimum values were recorded, particularly for the WD index. The usage of grid-based datasets may be deemed more reliable than the station-based approach as each pixel has its own hourly time series; however, these datasets are rarely available. The applicability of the station-based procedure in other alpine contexts is much more achievable and remains a valuable option, especially if resources could be devoted to the regionalization phase and performance metrics would return satisfying results, as in the present case study.

The verification of the suitability and completeness of the input data was carried out at several steps of the study and proved to be useful to: (i) plan an effective field survey to acquire new data in Valchiavenna, representative of the geomechanical context of the study area; (ii) determine the quality of the regionalized geomechanical and station-based climatic predictors implemented in the susceptibility models; (iii) to assess the IFFI rockfall inventory limitations and incompleteness in the complex mountainous environment of Valchiavenna; (iv) to demonstrate the existence of a strong inventory bias in the rockfall inventory of Aosta Valley and to plan a modelling configuration strategy to overcome this issue. Deepening the knowledge on local geomechanical and climatic conditions led to a process-oriented understanding of the susceptibility spatial patterns, improving the quality and readability of the susceptibility maps produced. To obtain reliable results, efforts should be made to derive regionalized maps of predictors specifically suited for the analysed landslide type and complete landslide inventories. As demonstrated through the analyses conducted in both study areas, the lack of these conditions can lead to erroneous outputs that could jeopardize an effective management of hazard and risk related to these phenomena, highlighting that the quantitative model performance alone is not sufficient to assess model reliability.

For both case studies, Generalized Additive Models (GAM) were used for rockfall susceptibility assessment; for the Valchiavenna case study, a Random Forest (RF) model was tested too. All models were validated through k-fold cross validation routines and their performance evaluated in terms of area under the receiver operating characteristic curve (AUROC). However, quantitative performance measures of the produced susceptibility models did not guarantee their geomorphological and physical plausibility. An in-depth critical analysis of the results in light of the specific environmental context was crucial in both case studies. Predictors' behaviour physical plausibility was thus verified through the analysis of the mathematical functions describing the predictors-susceptibility modelled relationships, i.e., Component Smoothing Functions for GAM and SHapely Additive exPlanations for RF. For Valchiavenna case study, the field acquisition and processing of geomechanical data was undoubtedly

time-consuming; with an AUROC improvement of 2%-3% respect to a basic topographic model it is legitimate to argue that the inclusion of such predictors was not worth the effort to derive them. However, results highlighted that including geomechanical predictors in a rockfall susceptibility model led to a redistribution of the high-susceptibility areas in plausible geomorphological contexts, such as in correspondence of active slope deformations and structural lineaments, otherwise not revealed by the topographic predictors alone. Moreover, an additional plausibility assessment of the output maps was conducted by means of secondary products such as variation maps and comparison with SAR data, useful to improve the readability and the communication of the results to potential users.

The consequences of the selection either of different susceptibility modelling algorithms or different model configuration setups on the outputs were analysed. For the Valchiavenna case study, Jv resulted as the most important geomechanical predictor both in the GAM (with a deviance explained of 7.5% and ranked 6th) and in the RF model (ranked 7th), with a rockfall susceptibility increase in correspondence of the most fractured rock masses. Wi and Keq were penalized (i.e., they had low influence on rockfall susceptibility) in the GAM model, whereas Keq showed an importance comparable to Jv in the RF model. In a complex Machine Learning model (RF), the SHAPs allowed the interpretation of predictors' behaviour, which demonstrated to be coherent with that shown in the GAM model. The discordance in predicting rockfall susceptibility between the GAM and the RF models was limited and varied from 13% to 8% of the total study area, using a 0.5 and 0.7 susceptibility cut-off threshold, respectively. Moreover, the demonstrated coherency of predictors behaviour amongst the two models enhanced the reliability and robustness of the predictors' role in defining susceptibility. For the Aosta Valley case study, rockfall susceptibility models, including both topographic, climatic and additional snow-related predictors (developed by Camera et al., 2021, from a SWE weekly gridded dataset) were developed and compared. Starting from a "Blind model", the susceptibility analysis was optimized through stepwise modifications of the model settings to handle issues related to inventory bias (through visibility masks), physical significance of climatic predictors (through CSF graphs analyses) and concurvity (through Principal Component Analysis). In the "Blind" models, ignoring inventory bias led to excellent model performances ($0.8 \leq AUROC \leq 0.90$) but physically implausible and distorted representation of rockfall susceptibility in terms of both predictors' behaviour (e.g., elevation, EWI, FT and cumulated number of snow melting events in a hydrological year - SWEep) and susceptibility spatial pattern. Most susceptible areas were relegated to the main valley bottoms, while high altitude areas far from roads and infrastructure were characterized by very low predicted susceptibility. The selection of non-rockfall points inside the defined "visibility mask" was a valuable approach to manage and reduce the inventory bias influence. The inclusion of climate predictors resulted in an improvement of the

susceptibility model performance (AUROC up to 3%) in comparison to a topographic-based model. Conversely, not-physically plausible climatic predictors were excluded from the models, even if their presence increased the model quantitative performance. The most important physically plausible climate predictors were EWI, WD, with a deviance explained varying from 5% to 10% each, followed by the maximum cumulated snow melting calculated on 32-day periods (SWEmax) with a deviance explained varying from 3% to 5%. The predictor SWEep (i.e., cumulated number of snow melting events in a hydrological year) was always represented by a physically distorted CSF, while FT was almost always penalized as its behaviour was masked by elevation. However, in presence of concurvity, predictors penalization may imply that other covariates have a stronger signal on the response variable, and not necessarily a missing relationship amongst the penalized variable and the investigated phenomena. When the climate and snow related predictors were inserted in the susceptibility model as principal components, concurvity was efficiently reduced. Nonetheless, despite PCA-related models embody the quality of parsimony, they may suffer from a less immediate interpretability.

Through the analyses conducted for the Aosta Valley case study, a process-related, potentially non-stationary configuration to the susceptibility models was offered. This is fundamental to implement dynamic susceptibility, hazard and risk analyses, which are necessary in light of the climate changes affecting mountain environments. The availability of a large historical rockfall inventory and an extensive, multi-variable meteorological dataset for the period 1990-2020 were crucial inputs for the analysis. Dated rockfall events, which are rarely available, and a good spatial-temporal resolution of meteorological data were fundamental to carry out this part of the project. For this reason, the efforts that some administrations (e.g., Aosta Valley Region) are putting in developing not only spatial- but also temporal-detailed inventories are extremely important. The temporal information is the one that allows to analyse the influence on landslides occurrence of changing boundary conditions, such as climatic processes, land use and vegetation.

Acknowledgments

For Valchiavenna Case Study:
I would like to acknowledge Comunità Montana Valchiavenna for the access to local facilities and field work permissions, and the Interreg V-A Italy-Switzerland Cooperation Program, A.M.AL.PI.2018 "Alpi in Movimento, Movimento nelle Alpi. Piuro 1618-2018"- id 594274 for SAR data provision.

For Aosta Valley Case Study:
I would like to acknowledge Centro Funzionale Regione Valle d'Aosta, ARPA Valle d'Aosta and the Regional Geological Office for meteorological and geological data provision and technical support.

I wish also to thank Professor Alexander Brenning and the GIScience Group from the Institute of Geography of the Friedrich Schiller Universität Jena, where part of the PhD project was carried out, for welcoming me in their group and helping me to improve my skills and knowledge on landslide susceptibility modelling.

I wish finally to thank the two reviewers of the thesis, Dr. Roberta Paranunzio and Prof. Renato Macciotta, for their valuable comments and constructive critics, which helped to improve the level of the work.

References

Abdi, H., Williams, L.J., 2010. Principal component analysis. WIREs Computational Statistics 2, 433–459. https://doi.org/10.1002/wics.101

Abedini, M., Ghasemian, B., Shirzadi, A., Bui, D.T., 2019. A comparative study of support vector machine and logistic model tree classifiers for shallow landslide susceptibility modeling. Environmental Earth Sciences 78, 560. https://doi.org/10.1007/s12665-019-8562-z

Abella, E.A.C., Westen, C.J.V., 2008. Qualitative landslide susceptibility assessment by multicriteria analysis: A case study from San Antonio del Sur, Guantánamo, Cuba. Geomorphology 94, 453–466. https://doi.org/10.1016/j.geomorph.2006.10.038

Adnan, M.S.G., Rahman, M.S., Ahmed, N., Ahmed, B., Rabbi, Md.F., Rahman, R.M., 2020. Improving Spatial Agreement in Machine Learning-Based Landslide Susceptibility Mapping. Remote Sensing 12. https://doi.org/10.3390/rs12203347

Agostini, A., Tofani, V., Nolesini, T., Gigli, G., Tanteri, L., Rosi, A., Cardellini, S., Casagli, N., 2014. A new appraisal of the Ancona landslide based on geotechnical investigations and stability modelling. Quarterly Journal of Engineering Geology and Hydrogeology 47, 29–43. https://doi.org/10.1144/qjegh2013-028

Akgun, A., 2012. A comparison of landslide susceptibility maps produced by logistic regression, multi-criteria decision, and likelihood ratio methods: a case study at İzmir, Turkey. Landslides 9, 93–106. https://doi.org/10.1007/s10346-011-0283-7

Allen, S., Huggel, C., 2013. Extremely warm temperatures as a potential cause of recent high mountain rockfall. Global and Planetary Change 107, 59–69. https://doi.org/10.1016/j.gloplacha.2013.04.007

Alqadhi, S., Mallick, J., Talukdar, S., Bindajam, A.A., Van Hong, N., Saha, T.K., 2021. Selecting optimal conditioning parameters for landslide susceptibility: an experimental research on Aqabat Al-Sulbat, Saudi Arabia. Environmental Science and Pollution Research. https://doi.org/10.1007/s11356-021-15886-z

Althuwaynee, O.F., Pradhan, B., Park, H.-J., Lee, J.H., 2014. A novel ensemble bivariate statistical evidential belief function with knowledge-based analytical hierarchy process and multivariate statistical logistic regression for landslide susceptibility mapping. CATENA 114, 21–36. https://doi.org/10.1016/j.catena.2013.10.011

Alvioli, M., Marchesini, I., Reichenbach, P., Rossi, M., Ardizzone, F., Fiorucci, F., Guzzetti, F., 2016. Automatic delineation of geomorphological slope units with \textttr.slopeunits v1.0 and their optimization for landslide susceptibility modeling. Geoscientific Model Development 9, 3975–3991. https://doi.org/10.5194/gmd-9-3975-2016

Amato, G., Eisank, C., Castro-Camilo, D., Lombardo, L., 2019. Accounting for covariate distributions in slope-unit-based landslide susceptibility models. A case study in the alpine environment. Engineering Geology 260, 105237. https://doi.org/10.1016/j.enggeo.2019.105237

Amato, G., Fiorucci, M., Martino, S., Lombardo, L., and Palombi, L., 2021. Earthquake-triggered landslide susceptibility in Italy by means of Artificial Neural Network. EarthArXiv. https://doi.org/10.31223/X59W39

Ambrosi, C., Crosta, G.B., 2006. Large sackung along major tectonic features in the Central Italian Alps. Engineering Geology 83, 183–200. https://doi.org/10.1016/j.enggeo.2005.06.031

Ambrosi, C., Crosta, G.B., 2011. Valley shape influence on deformation mechanisms of rock slopes. In: Jaboyedoff, M. (Ed.), Slope Tectonics. Geological Society, London, pp. 215–233.

Amitrano, D., Arattano, M., Chiarle, M., Mortara, G., Occhiena, C., Pirulli, M., Scavia, C., 2010. Microseismic activity analysis for the study of the rupture mechanisms in unstable rock masses. Natural Hazards and Earth System Sciences 10, 831–841. https://doi.org/10.5194/nhess-10-831-2010

Amodio, S., Aria, M., D'Ambrosio, A., 2014. On concurvity in nonlinear and nonparametric regression models. Statistica 74, 85–98.

Antonielli, B., Mazzanti, P., Rocca, A., Bozzano, F., Dei Cas, L., 2019. A-DInSAR Performance for Updating Landslide Inventory in Mountain Areas: An Example from Lombardy Region (Italy). Geosciences 9. https://doi.org/10.3390/geosciences9090364

Arabameri, A., Chen, W., Loche, M., Zhao, X., Li, Y., Lombardo, L., Cerda, A., Pradhan, B., Bui, D.T., 2020. Comparison of machine learning models for gully erosion susceptibility mapping. Geoscience Frontiers 11, 1609–1620. https://doi.org/10.1016/j.gsf.2019.11.009

Ardennes, Belgium. Natural Hazards and Earth System Sciences 9, 507–521. https://doi.org/10.5194/nhess-9-507-2009

Ardizzone, F., Cardinali, M., Carrara, A., Guzzetti, F., Reichenbach, P., 2002. Impact of mapping errors on the reliability of landslide hazard maps. Natural Hazards and Earth System Sciences 2, 3–14. https://doi.org/10.5194/nhess-2-3-2002

Argand, E. 1911. Les nappes des recouvrement des Alpes Pennines et leur prolongements structuraux. Mat. Carte géol. Suisse, 31, 1–26.

Attorre, F., Alfo', M., De Sanctis, M., Francesconi, F., Bruno, F., 2007. Comparison of interpolation methods for mapping climatic and bioclimatic variables at regional scale. Int. J. Climatol. 27, 1825–1843. https://doi.org/10.1002/joc.1495

Auer, I., Böhm, R., Jurkovic, A., Lipa, W., Orlik, A., Potzmann, R., Schöner, W., Ungersböck, M., Matulla, C., Briffa, K., Jones, P., Efthymiadis, D., Brunetti, M., Nanni, T., Maugeri, M., Mercalli, L., Mestre, O., Moisselin, J.-M., Begert, M., Müller-Westermeier, G., Kveton, V., Bochnicek, O., Stastny, P., Lapin, M., Szalai, S., Szentimrey, T., Cegnar, T., Dolinar, M., Gajic-Capka, M., Zaninovic, K., Majstorovic, Z., Nieplova, E., 2007. HISTALP—historical instrumental climatological surface time series of the Greater Alpine Region. International Journal of Climatology 27, 17–46. https://doi.org/10.1002/joc.1377

Augustinus, P.C., 1992. The influence of rock mass strength on glacial valley cross-section morphometry: a case study from the Southern Alps, New Zealand. EarthSurface Processes and Landforms 17, 39–51.

Augustinus, P.C., 1995. Glacial valley cross-profile development: the influence of in situ rock stress and rock mass strength, with examples from the Southern Alps, New Zealand. Geomorphology 14, 87–97.

Ba, Q., Chen, Y., Deng, S., Yang, J., Li, H., 2018. A comparison of slope units and grid cells as mapping units for landslide susceptibility assessment. Earth Science Informatics 11, 373–388. https://doi.org/10.1007/s12145-018-0335-9

Bagalwa, R.M., Chartin, C., Baumgartner, S., Mercier, S., Syauswa, M., Samba, V.C., Zabona, M.T., Karume, K., Cizungu, N.L., Barthel, M., Doetterl, S., Six, J., Boeckx, P., Oost, K.V., 2021. Spatial and seasonal patterns of rainfall erosivity in the Lake Kivu region: Insights from a meteorological

observatory network. Progress in Physical Geography: Earth and Environment 0, 03091333211001793. https://doi.org/10.1177/03091333211001793

Baghbanan, A., Jing, L., 2007. Hydraulic properties of fractured rock masses with correlated fracture length and aperture. International Journal of Rock Mechanics and Mining Sciences 44, 704–719. https://doi.org/10.1016/j.ijrmms.2006.11.001

Baghbanan, A., Jing, L., 2008. Stress effects on permeability in a fractured rock mass with correlated fracture length and aperture. International Journal of Rock Mechanics and Mining Sciences 45, 1320–1334. https://doi.org/10.1016/j.ijrmms.2008.01.015

Bajni, G., Camera, C.A.S., Apuani, T., 2021a. Deciphering meteorological influencing factors for Alpine rockfalls: a case study in Aosta Valley. Landslides 18, 3279–3298. https://doi.org/10.1007/s10346-021-01697-3

Bajni, G., Camera, C.A.S., Brenning, A., Apuani, T., 2021b. Rock mass geomechanical properties to improve rockfall susceptibility assessment: a case study in Valchiavenna (SO). IOP Conference Series: Earth and Environ. Sci., vol. 833 012180. Mechanics and Rock Engineering, from Theory to Practice, 20-25 September 2021, Turin, Italy https://doi.org/10.1088/1755-1315/833/1/012180

Bakun-Mazor, D., Hatzor, Y.H., Glaser, S.D., Santamarina, J.C., 2013. Thermally vs. seismically induced block displacements in Masada rock slopes. International Journal of Rock Mechanics and Mining Sciences 61, 196–211. https://doi.org/10.1016/j.ijrmms.2013.03.005

Ballantyne, C.K., 2002. Paraglacial geomorphology. Quaternary Science Reviews 21, 1935–2017. https://doi.org/10.1016/S0277-3791(02)00005-7

Ballantyne, C.K., Stone, J.O., 2013. Timing and periodicity of paraglacial rock-slope failures in the Scottish Highlands. Geomorphology 186, 150–161. https://doi.org/10.1016/j.geomorph.2012.12.030

Ballèvre, M., Kienast, J. R., and Vuichard, J. P., 1986. La “nappe de la Dent Blanche” (Alpes occidentales): deux unites austroalpines indépéndantes. Eclogae Geologicae Helvetiae, 79, 57–74.

Ballèvre, M., Manzotti, P., Le Bayon, B., Real, C., 2015. From maps and cross-sections to kinematic models: The impact of metamorphic petrology. Rendiconti Online della Società Geologica Italiana 37, 5–8. https://doi.org/10.3301/ROL.2015.181

Ban, N., Rajczak, J., Schmidli, J., Schär, C., 2020. Analysis of Alpine precipitation extremes using generalized extreme value theory in convection-resolving climate simulations. Climate Dynamics 55, 61–75. https://doi.org/10.1007/s00382-018-4339-4

Barca, E., Passarella, G., Vurro, M., Morea, A., 2015. MSANOS: Data-Driven, Multi-Approach Software for Optimal Redesign of Environmental Monitoring Networks. Water Resour Manage 29, 619–644. https://doi.org/10.1007/s11269-014-0859-9

Barton, N., 2004a. The theory behind high pressure grouting—Part 1. Tunnels & Tunnelling International, September, pp. 28–30.

Barton, N., 2004b. The theory behind high pressure grouting—Part 2. Tunnels & Tunnelling International, October, pp. 33–35.

Barton, N., Lien, R. and Lunde, J., 1974. Engineering Classification of Rock Masses for the Design of Tunnel Support. Rock Mechanics, 6, 189-236. https://doi.org/10.1007/BF01239496

Baudin T.H. and Marquer D., 1993. Metamorphism and deformation in the Tambò nappe (Swiss Central Alps): evolution of the phengite substitution during Alpine deformation. Schweizer Mineralogische und Petrographische Mitteilungen, Vol.73, 285-299

Baudin T.H., Marquer D., Barfety J.C., Kerckhove C. & Persoz F., 1995. A new stratigraphical interpretation of the mesozoic cover of the Tambò and Suretta nappes: Evidence for early thin-skinned tectonics (Swiss Central Alps). Comptes Rendus de l'Academie des Sciences Paris, Vol.321 (5), 401-408

Behling, R., Roessner, S., Kaufmann, H., Kleinschmit, B., 2014. Automated Spatiotemporal Landslide Mapping over Large Areas Using RapidEye Time Series Data. Remote Sensing 6, 8026–8055. https://doi.org/10.3390/rs6098026

Bell, R., Petschko, H., Röhrs, M., Dix, A., 2012. Assessment of landslide age, landslide persistence and human impact using airborne laser scanning digital terrain models. Geografiska Annaler: Series A, Physical Geography 94, 135–156. https://doi.org/10.1111/j.1468-0459.2012.00454.x

Ben-Ameur, W., 2004. Computing the Initial Temperature of Simulated Annealing. Computational Optimization and Applications 29, 369–385. https://doi.org/10.1023/B:COAP.0000044187.23143.bd

Bender, E., Lehning, M., Fiddes, J., 2020. Changes in Climatology, Snow Cover, and Ground Temperatures at High Alpine Locations. Frontiers in Earth Science 8, 100. https://doi.org/10.3389/feart.2020.00100

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

Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D., Andreassian, V., 2013. Characterising performance of environmental models. Environmental Modelling & Software 40, 1–20. https://doi.org/10.1016/j.envsoft.2012.09.011

Berardino, P., Fornaro, G., Lanari, R., Sansosti, E., 2002. A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms. IEEE Transactions on Geoscience and Remote Sensing 40, 2375–2383. https://doi.org/10.1109/TGRS.2002.803792

Berger A, Rosenberg C, Schmid SM. 1996. Ascent, emplacement and exhumation of the Bergell pluton within the southern steep belt of the Central Alps. Schweizerische Mineralogische Petrographische Mitteilungen 76, 357–382.

Bianchini, S., Herrera, G., Mateos, R.M., Notti, D., Garcia, I., Mora, O., Moretti, S., 2013. Landslide Activity Maps Generation by Means of Persistent Scatterer Interferometry. Remote Sensing 5, 6198–6222. https://doi.org/10.3390/rs5126198

Bieniawski, Z.T., 1973. Engineering Classification of Jointed Rock Masses. Transaction of the South African Institution of Civil Engineers, 15, 335-344.

Bigi G., Castellarin A., Coli M., Dal Piaz G.V., Sartori R., Scandone P., Vai G.B., 1990. Structural Model of Italy scale 1:500.000, sheet 1. C.N.R., Progetto Finalizzato Geodinamica, SELCA Firenze.

Bistacchi A, Dal Piaz G.V., Massironi M., Zattin M. And Balestrieri M.L., 2001 -The Aosta-Ranzola extensional fault system and Oligocene-Present evolution of the north-western Alpine nappe stack. Int. J. Earth Sci., v. 90, pp. 654-667.

Bodin, X., Schoeneich, P., Deline, P., Ravanel, L., Magnin, F., Krysiecki, J.-M., Echelard, T., 2015. Mountain permafrost and associated geomorphological processes: recent changes in the French Alps. Revue De Geographie Alpine-journal of Alpine Research.

Boehmke, B., and Greenwell, B., 2019. Hands-On Machine Learning with R (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9780367816377

Bonham-Carter, G.F., 1994. Geographic Information Systems for Geoscientists: Modelling with GIS. Pergamon, Ottawa (398 pp).

Booth, A.M., Roering, J.J., Perron, J.T., 2009. Automated landslide mapping using spectral analysis and high-resolution topographic data: Puget Sound lowlands, Washington, and Portland Hills, Oregon. Geomorphology 109, 132–147. https://doi.org/10.1016/j.geomorph.2009.02.027

Bordoni, M., Vivaldi, V., Lucchelli, L., Ciabatta, L., Brocca, L., Galve, J.P., Meisina, C., 2021. Development of a data-driven model for spatial and temporal shallow landslide probability of occurrence at catchment scale. Landslides 18, 1209–1229. https://doi.org/10.1007/s10346-020-01592-3

Bornaetxea, T., Rossi, M., Marchesini, I., Alvioli, M., 2018. Effective surveyed area and its role in statistical landslide susceptibility assessments. Natural Hazards and Earth System Sciences 18, 2455–2469. https://doi.org/10.5194/nhess-18-2455-2018

Boulesteix, A.-L., Janitza, S., Kruppa, J., König, I.R., 2012. Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. WIREs Data Mining and Knowledge Discovery 2, 493–507. https://doi.org/10.1002/widm.1072

Bousquet R., Schmid S.M., Zeilinger G., Oberhänsli R., Rosenberg C. L., Molli G., Robert C., Wiederkehr M., Rossi P.,2012. Tectonic framework of the Alps. CCGM/CGMW

Brabb, E.E., 1984. Innovative approaches to landslide hazard mapping. In: Proc. 4th Int. Symp. Landslides, Toronto. 1. pp. 307–324.

Brace, W.F., 1980. Permeability of crystalline and argillaceous rocks. International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts 17, 241–251. https://doi.org/10.1016/0148-9062(80)90807-4

Breiman, L., 2001. Random Forests. Machine Learning, 45, 5–32. https://doi.org/10.1201/9780367816377-11

Brenning, A., 2005. Spatial prediction models for landslide hazards: review, comparison and evaluation. Natural Hazards and Earth System Sciences 5, 853–862. https://doi.org/10.5194/nhess-5-853-2005

Brenning, A., 2009. Benchmarking classifiers to optimally integrate terrain analysis and multispectral remote sensing in automatic rock glacier detection. Remote Sensing of Environment 113, 239–247. https://doi.org/10.1016/j.rse.2008.09.005

Brenning, A., 2012b. Improved spatial analysis and prediction of landslide susceptibility: Practical recommendations. Landslides and Engineered Slopes, Protecting Society through Improved Understanding, edited by: Eberhardt, E., Froese, C., Turner, AK, and Leroueil, S., Taylor & Francis, Banff, Alberta, Canada 789–795.

Brenning, A., Grasser, M., Friend, D.A., 2007. Statistical estimation and generalized additive modeling of rock glacier distribution in the San Juan Mountains, Colorado, United States. Journal of Geophysical Research: Earth Surface 112. https://doi.org/10.1029/2006JF000528

Brenning, A.; Bangs, D.; Becker, M., 2018. RSAGA: SAGA Geoprocessing and Terrain Analysis. R Package Version 1.3.0. Available online: https://CRAN.R-project.org/package=RSAGA

Brenning, A.2012a. Spatial cross-validation and bootstrap for the assessment of prediction rules in remote sensing: the R package “sperrorest”, IEEE Int. Geosci. Remote Se., 23, 5372–5375.

Brenning, A., 2008. Statistical geocomputing combining R and SAGA: The example of landslide susceptibility analysis with generalized additive models. In: J. Böhner, T. Blaschke & L. Montanarella (eds.), SAGA – Seconds Out (Hamburger Beiträge zur Physischen Geographie und Landschaftsökologie, vol. 19), p. 23-32.

Brideau, M.-A., Pedrazzini, A., Stead, D., Froese, C., Jaboyedoff, M., van Zeyl, D., 2011. Three-dimensional slope stability analysis of South Peak, Crowsnest Pass, Alberta, Canada. Landslides 8, 139–158. https://doi.org/10.1007/s10346-010-0242-8

Broeckx, J., Vanmaercke, M., Duchateau, R., Poesen, J., 2018. A data-based landslide susceptibility map of Africa. Earth-Science Reviews 185, 102–121. https://doi.org/10.1016/j.earscirev.2018.05.002

Brus, D.J., 2019. Sampling for digital soil mapping: A tutorial supported by R scripts. Geoderma 338, 464–480. https://doi.org/10.1016/j.geoderma.2018.07.036

Brus, D.J., de Gruijter, J.J., 1997. Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil (with discussion). Geoderma 80, 1–44. https://doi.org/10.1016/S0016-7061(97)00072-4

Brus, D.J., Heuvelink, G.B.M., 2007. Optimization of sample patterns for universal kriging of environmental variables. Geoderma 138, 86–95. https://doi.org/10.1016/j.geoderma.2006.10.016

Bueechi, E., Klimeš, J., Frey, H., Huggel, C., Strozzi, T., Cochachin, A., 2019. Regional-scale landslide susceptibility modelling in the Cordillera Blanca, Peru—a comparison of different approaches. Landslides 16, 395–407. https://doi.org/10.1007/s10346-018-1090-1

Bunce, C.M., Cruden, D.M., Morgenstern, N.R., 1997. Assessment of the hazard from rock fall on a highway. Can. Geotech. J. 34, 344–356. https://doi.org/10.1139/t97-009

Burrough, P. A. and McDonnell, R. A., 1998. Principles of Geographical Information Systems. Oxford University Press, Oxford.

Cama, M., Conoscenti, C., Lombardo, L., Rotigliano, E., 2016. Exploring relationships between grid cell size and accuracy for debris-flow susceptibility models: a test in the Giampilieri catchment (Sicily, Italy). Environmental Earth Sciences 75, 238. https://doi.org/10.1007/s12665-015-5047-6

Cama, M., Lombardo, L., Conoscenti, C., Agnesi, V., Rotigliano, E., 2015. Predicting storm-triggered debris flow events: application to the 2009 Ionian Peloritan disaster (Sicily, Italy). Nat. Hazards Earth Syst. Sci. 15, 1785–1806. https://doi.org/10.5194/nhess-15-1785-2015

Camera, C., Bruggeman, A., Hadjinicolaou, P., Michaelides, S., Lange, M.A., 2017b. Evaluation of a spatial rainfall generator for generating high resolution precipitation projections over orographically complex terrain. Stoch. Environ. Res. Risk. Assess. 31, 757-773. https://doi.org/10.1007/s00477-016-1239-1

Camera, C.A.S., Bajni, G., Corno, I., Raffa, M., Stevenazzi, S., Apuani, T., 2021. Introducing intense rainfall and snowmelt variables to implement a process-related non-stationary shallow landslide susceptibility analysis. Science of The Total Environment 147360. https://doi.org/10.1016/j.scitotenv.2021.147360

Camera, C.A.S., Zomeni, Z., Noller, J.S., Zissimos, A.M., Christoforou, I.C., Bruggeman, A., 2017a. A high resolution map of soil types and physical properties for Cyprus: A digital soil mapping optimization. Geoderma 285, 35–49. https://doi.org/10.1016/j.geoderma.2016.09.019

Can, R., Kocaman, S., Gokceoglu, C., 2021. A Comprehensive Assessment of XGBoost Algorithm for Landslide Susceptibility Mapping in the Upper Basin of Ataturk Dam, Turkey. Applied Sciences 11. https://doi.org/10.3390/app11114993

Caniani, D., Pascale, S., Sdao, F., Sole, A., 2008. Neural networks and landslide susceptibility: a case study of the urban area of Potenza. Natural Hazards 45, 55–72. https://doi.org/10.1007/s11069-007-9169-3

Carlà, T., Nolesini, T., Solari, L., Rivolta, C., Dei Cas, L., Casagli, N., 2019. Rockfall forecasting and risk management along a major transportation corridor in the Alps through ground-based radar interferometry. Landslides 16, 1425–1435. https://doi.org/10.1007/s10346-019-01190-y

Carrara A., Cardinali M., Guzzetti F., Reichenbach P., 1995. GIS Technology in Mapping Landslide Hazard. In: Carrara A., Guzzetti F. (eds) Geographical Information Systems in Assessing Natural Hazards. Advances in Natural and Technological Hazards Research, vol 5. Springer, Dordrecht. https://doi.org/10.1007/978-94-015-8404-3_8

Carrara, A., 1983. Multivariate models for landslide hazard evaluation. Journal of the International Association for Mathematical Geology 15, 403–426. https://doi.org/10.1007/BF01031290

Carrara, A., Cardinali, M., Detti, R., Guzzetti, F., Pasqui, V., Reichenbach, P., 1991. GIS techniques and statistical models in evaluating landslide hazard. Earth Surface Processes and Landforms 16 (5), 427–445.

Cascini, L., Fornaro, G., Peduto, D., 2010. Advanced low- and full-resolution DInSAR map generation for slow-moving landslide analysis at different scales. Engineering Geology 112, 29–42. https://doi.org/10.1016/j.enggeo.2010.01.003

Catani, F., 2021. Landslide detection by deep learning of non-nadiral and crowdsourced optical images. Landslides 18, 1025–1044. https://doi.org/10.1007/s10346-020-01513-4

Catani, F., Lagomarsino, D., Segoni, S., Tofani, V., 2013. Landslide susceptibility estimation by random forests technique: sensitivity and scaling issues. Natural Hazards and Earth System Sciences 13, 2815–2831. https://doi.org/10.5194/nhess-13-2815-2013

Ceryan, S., Tudes, S., Ceryan, N., 2008. Influence of weathering on the engineering properties of Harsit granitic rocks (NE Turkey). Bulletin of Engineering Geology and the Environment 67, 97–104. https://doi.org/10.1007/s10064-007-0115-0

Chacón, J., Irigaray, C., Fernández, T., El Hamdouni, R., 2006. Engineering geology maps: landslides and geographical information systems. Bulletin of Engineering Geology and the Environment 65, 341–411. https://doi.org/10.1007/s10064-006-0064-z

Chang, K.-T., Merghadi, A., Yunus, A.P., Pham, B.T., Dou, J., 2019. Evaluating scale effects of topographic variables in landslide susceptibility models using GIS-based machine learning techniques. Scientific Reports 9, 12296. https://doi.org/10.1038/s41598-019-48773-2

Chen, B., Pan, Y., Wang, J., Fu, Z., Zeng, Z., Zhou, Y., Zhang, Y., 2013. Even sampling designs generation by efficient spatial simulated annealing. Mathematical and Computer Modelling 58, 670–676. https://doi.org/10.1016/j.mcm.2011.10.035

Chen, G., Hay, G.J., Carvalho, L.M.T., Wulder, M.A., 2012. Object-based change detection. International Journal of Remote Sensing 33, 4434–4457. https://doi.org/10.1080/01431161.2011.648285

Chen, L., Guo, Z., Yin, K., Shrestha, D.P., Jin, S., 2019. The influence of land use and land cover change on landslide susceptibility: a case study in Zhushan Town, Xuan'en County (Hubei, China). Natural Hazards and Earth System Sciences 19, 2207–2228. https://doi.org/10.5194/nhess-19-2207-2019

Chen, W., Chen, Y., Tsangaratos, P., Ilia, I., Wang, X., 2020. Combining Evolutionary Algorithms and Machine Learning Models in Landslide Susceptibility Assessments. Remote Sensing 12. https://doi.org/10.3390/rs12233854

Chen, W., Li, Y., 2020. GIS-based evaluation of landslide susceptibility using hybrid computational intelligence models. CATENA 195, 104777. https://doi.org/10.1016/j.catena.2020.104777

Chen, W., Peng, J., Hong, H., Shahabi, H., Pradhan, B., Liu, J., Zhu, A.-X., Pei, X., Duan, Z., 2018. Landslide susceptibility modelling using GIS-based machine learning techniques for Chongren County, Jiangxi Province, China. Science of The Total Environment 626, 1121–1135. https://doi.org/10.1016/j.scitotenv.2018.01.124

Chen, W., Pourghasemi, H.R., Panahi, M., Kornejady, A., Wang, J., Xie, X., Cao, S., 2017. Spatial prediction of landslide susceptibility using an adaptive neuro-fuzzy inference system combined with frequency ratio, generalized additive model, and support vector machine techniques. Geomorphology 297, 69–85. https://doi.org/10.1016/j.geomorph.2017.09.007

Chen, W., Sun, Z., Han, J., 2019. Landslide Susceptibility Modeling Using Integrated Ensemble Weights of Evidence with Logistic Regression and Random Forest Models. Applied Sciences 9. https://doi.org/10.3390/app9010171

Chen, Z., Ye, F., Fu, W., Ke, Y., Hong, H., 2020. The influence of DEM spatial resolution on landslide susceptibility mapping in the Baxie River basin, NW China. Natural Hazards 101, 853–877. https://doi.org/10.1007/s11069-020-03899-9

Chiarle, M., Geertsema, M., Mortara, G., Clague, J.J., 2021. Relations between climate change and mass movement: Perspectives from the Canadian Cordillera and the European Alps. Global and Planetary Change 202, 103499. https://doi.org/10.1016/j.gloplacha.2021.103499

Choi, J., Oh, H.-J., Lee, H.-J., Lee, C., Lee, S., 2012. Combining landslide susceptibility maps obtained from frequency ratio, logistic regression, and artificial neural network models using ASTER images and GIS. Engineering Geology 124, 12–23. https://doi.org/10.1016/j.enggeo.2011.09.011

Chung, C.-J.F., Fabbri, A.G., 2003. Validation of Spatial Prediction Models for Landslide Hazard Mapping. Natural Hazards 30, 451–472. https://doi.org/10.1023/B:NHAZ.0000007172.62651.2b

Ciampalini, A., Raspini, F., Lagomarsino, D., Catani, F., Casagli, N., 2016a. Landslide susceptibility map refinement using PSInSAR data. Remote Sensing of Environment 184, 302–315. https://doi.org/10.1016/j.rse.2016.07.018

Ciampalini, A., Raspini, F., Lagomarsino, D., Catani, F., Casagli, N., 2016b. Landslide susceptibility map refinement using PSInSAR data. Remote Sensing of Environment 184, 302–315. https://doi.org/10.1016/j.rse.2016.07.018

Cignetti, M., Godone, D., Bertolo, D., Paganone, M., Thuegaz, P., Giordan, D., 2020. Rockfall susceptibility along the regional road network of Aosta Valley Region (northwestern Italy). Journal of Maps 0, 1–11. https://doi.org/10.1080/17445647.2020.1850534

Citrini, A., Camera, C., Beretta, G.P., 2020. Nossana Spring (Northern Italy) under Climate Change: Projections of Future Discharge Rates and Water Availability. Water 12. https://doi.org/10.3390/w12020387

Clerici, A., Perego, S., Tellini, C., Vescovi, P., 2002. A procedure for landslide susceptibility zonation by the conditional analysis method. Geomorphology 48, 349–364. https://doi.org/10.1016/S0169-555X(02)00079-X

Cochran, W.G., 1977. Sampling Techniques, third ed. John Wiley & Sons, USA.

Coe, J.A., Harp, E.L., 2007. Influence of tectonic folding on rockfall susceptibility, American Fork Canyon, Utah, USA. Natural Hazards and Earth System Sciences 7, 1–14. https://doi.org/10.5194/nhess-7-1-2007

Colesanti, C., Wasowski, J., 2006. Investigating landslides with space-borne Synthetic Aperture Radar (SAR) interferometry. Engineering Geology 88, 173–199. https://doi.org/10.1016/j.enggeo.2006.09.013

Coli, N., Pranzini, G., Alfi, A., Boerio, V., 2008. Evaluation of rock-mass permeability tensor and prediction of tunnel inflows by means of geostructural surveys and finite element seepage analysis. Engineering Geology 101, 174–184. https://doi.org/10.1016/j.enggeo.2008.05.002

Collins, B.D., Stock, G.M., 2016. Rockfall triggering by cyclic thermal stressing of exfoliation fractures. Nature Geoscience 9, 395–400. https://doi.org/10.1038/ngeo2686

Collins, B.D., Stock, G.M., Eppes, M.-C., Lewis, S.W., Corbett, S.C., Smith, J.B., 2018. Thermal influences on spontaneous rock dome exfoliation. Nature Communications 9, 762. https://doi.org/10.1038/s41467-017-02728-1

Conoscenti, C., Rotigliano, E., Cama, M., Caraballo-Arias, N.A., Lombardo, L., Agnesi, V., 2016. Exploring the effect of absence selection on landslide susceptibility models: A case study in Sicily, Italy. Geomorphology 261, 222–235. https://doi.org/10.1016/j.geomorph.2016.03.006

Corò, D., Galgaro, A., Fontana, A., Carton, A., 2015. A regional rockfall database: the Eastern Alps test site. Environmental Earth Sciences 74, 1731–1742. https://doi.org/10.1007/s12665-015-4181-5

Corominas, J., van Westen, C., Frattini, P., Cascini, L., Malet, J.-P., Fotopoulou, S., Catani, F., Van Den Eeckhaut, M., Mavrouli, O., Agliardi, F., Pitilakis, K., Winter, M.G., Pastor, M., Ferlisi, S., Tofani, V., Hervás, J., Smith, J.T., 2014. Recommendations for the quantitative analysis of landslide risk. Bulletin of Engineering Geology and the Environment 73, 209–263. https://doi.org/10.1007/s10064-013-0538-8

Crippa, C., Franzosi, F., Zonca, M., Manconi, A., Crosta, G.B., Dei Cas, L., Agliardi, F., 2020. Unraveling Spatial and Temporal Heterogeneities of Very Slow Rock-Slope Deformations with Targeted DInSAR Analyses. Remote Sensing 12. https://doi.org/10.3390/rs12081329

Crosta, G.B., di Prisco, C., Frattini, P., Frigerio, G., Castellanza, R., Agliardi, F., 2014. Chasing a complete understanding of the triggering mechanisms of a large rapidly evolving rockslide. Landslides 11, 747–764. https://doi.org/10.1007/s10346-013-0433-1

Crosta, G.B., Frattini, P., Agliardi, F., 2013. Deep seated gravitational slope deformations in the European Alps. Tectonophysics 605, 13–33. https://doi.org/10.1016/j.tecto.2013.04.028

Crozier, M. J., & Glade, T., 2005. Landslide Hazard and Risk: Issues, Concepts and Approach. In T. Glade, M. G. Anderson, & M. J. Crozier, Eds., Landslide Risk Assessment (pp. 1-40). New York: John Wiley. https://doi.org/10.1002/9780470012659.ch1

Crozier, M.J., 2010. Deciphering the effect of climate change on landslide activity: A review. Geomorphology 124, 260–267. https://doi.org/10.1016/j.geomorph.2010.04.009

Cruden D.M., Varnes D.J., 1996. Landslide types and processes. In: Turner AK, Schuster RL (eds) Landslides: investigation and mitigation, transportation research board special report 247. National Research Council, USA, pp 36–75

Cruden, D.M., 2003. The shapes of cold, high mountains in sedimentary rocks. Geomorphology 55, 249–261. https://doi.org/10.1016/S0169-555X(03)00143-0

D'Amato, J., Hantz, D., Guerin, A., Jaboyedoff, M., Baillet, L., Mariscal, A., 2016. Influence of meteorological factors on rockfall occurrence in a middle mountain limestone cliff. Natural Hazards and Earth System Sciences 16, 719–735. https://doi.org/10.5194/nhess-16-719-2016

Dal Piaz, G., Cortiana, G., Del Moro, A., Martin, S., Pennacchioni, G., Tartarotti, P., 2001. Tertiary age and paleostructural inferences of the eclogitic imprint in the Austroalpine outliers and Zermatt–Saas ophiolite, western Alps. International Journal of Earth Sciences 90, 668–684. https://doi.org/10.1007/s005310000177

de Gruijter, J., Brus, D.J., Bierkens, M.F., Knotters, M., 2006. Sampling for natural resource monitoring. Springer.

de Gruijter, J.J., McBratney, A.B., Taylor, J., 2010. Sampling for High Resolution Soil Mapping. Springer Netherlands, Sydney, Australia, pp. 3–14 chapter 1.

De Vilder, S.J., Rosser, N.J., Brain, M.J., 2017. Forensic analysis of rockfall scars. Geomorphology 295, 202–214. https://doi.org/10.1016/j.geomorph.2017.07.005

Delonca, A., Gunzburger, Y., Verdel, T., 2014. Statistical correlation between meteorological and rockfall databases. Natural Hazards and Earth System Sciences 14, 1953–1964. https://doi.org/10.5194/nhess-14-1953-2014

DeWalle DR, Rango A (2008) Principles of snow hydrology. Cambridge University Press, New York, 410 pp. https://doi.org/10.1657/1938-4246-41.4.523

Di Napoli, M., Carotenuto, F., Cevasco, A., Confuorto, P., Di Martire, D., Firpo, M., Pepe, G., Raso, E., Calcaterra, D., 2020. Machine learning ensemble modelling as a tool to improve landslide susceptibility mapping reliability. Landslides 17, 1897–1914. https://doi.org/10.1007/s10346-020-01392-9

Díaz-Uriarte, R., Alvarez de Andrés, S., 2006. Gene selection and classification of microarray data using random forest. BMC Bioinformatics 7, 3. https://doi.org/10.1186/1471-2105-7-3

Dikshit, A., Sarkar, R., Pradhan, B., Jena, R., Drukpa, D., Alamri, A.M., 2020. Temporal Probability Assessment and Its Use in Landslide Susceptibility Mapping for Eastern Bhutan. Water 12. https://doi.org/10.3390/w12010267

Dorren, L.K.A., 2003. A review of rockfall mechanics and modelling approaches. Progress in Physical Geography: Earth and Environment 27, 69–87. https://doi.org/10.1191/0309133303pp359ra

Dou, J., Yunus, A.P., Bui, D.T., Merghadi, A., Sahana, M., Zhu, Z., Chen, C.-W., Han, Z., Pham, B.T., 2020. Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan. Landslides 17, 641–658. https://doi.org/10.1007/s10346-019-01286-5

Draebing, D., 2021. Identification of rock and fracture kinematics in high alpine rockwalls under the influence of elevation. Earth Surface Dynamics 9, 977–994. https://doi.org/10.5194/esurf-9-977-2021

Draebing, D., Krautblatter, M., 2019. The Efficacy of Frost Weathering Processes in Alpine Rockwalls. Geophysical Research Letters 46, 6516–6524. https://doi.org/10.1029/2019GL081981

Drăguţ, L., Blaschke, T., 2006. Automated classification of landform elements using object-based image analysis. Geomorphology 81, 330–344. https://doi.org/10.1016/j.geomorph.2006.04.013

Drăguţ, L., Eisank, C., 2012. Automated object-based classification of topography from SRTM data. Geomorphology 141–142, 21–33. https://doi.org/10.1016/j.geomorph.2011.12.001

Du, G., Zhang, Y., Iqbal, J., Yang, Z., Yao, X., 2017. Landslide susceptibility mapping using an integrated model of information value method and logistic regression in the Bailongjiang watershed, Gansu Province, China. J. Mt. Sci. 14, 249–268. https://doi.org/10.1007/s11629-016-4126-9

Du, L., You, X., Li, K., Meng, L., Cheng, G., Xiong, L., Wang, G., 2019. Multi-modal deep learning for landform recognition. ISPRS Journal of Photogrammetry and Remote Sensing 158, 63–75. https://doi.org/10.1016/j.isprsjprs.2019.09.018

Du, M., Liu, N., Hu, X. 2019. Techniques for interpretable machine learning. Communications of the ACM, 63(1), 68-77.

Duan, Q., Yang, X., Chen, J., 2017. Hydraulic properties of a low permeable rupture zone on the Yingxiu-Beichuan Fault activated during the Wenchuan earthquake, China: Implications for fluid conduction, fault sealing, and dynamic weakening mechanisms. Tectonophysics 721, 123–142. https://doi.org/10.1016/j.tecto.2017.10.002

Dunham, L., Wartman, J., Olsen, M.J., O'Banion, M., Cunningham, K., 2017. Rockfall Activity Index (RAI): A lidar-derived, morphology-based method for hazard assessment. Engineering Geology 221, 184–192. https://doi.org/10.1016/j.enggeo.2017.03.009

Duvillard, P.-A., Ravanel, L., Deline, P., 2015. Risk assessment of infrastructure destabilisation due to global warming in the high French Alps. Journal of Alpine Research| Revue de géographie alpine.

Eberhardt, E., Thuro, K., Luginbuehl, M., 2005. Slope instability mechanisms in dipping interbedded conglomerates and weathered marls—the 1999 Rufi landslide, Switzerland. Engineering Geology 77, 35–56. https://doi.org/10.1016/j.enggeo.2004.08.004

Eeckhaut, M.V.D., Kerle, N., Poesen, J., Hervás, J., 2012. Object-oriented identification of forested landslides with derivatives of single pulse LiDAR data. Geomorphology 173–174, 30–42. https://doi.org/10.1016/j.geomorph.2012.05.024

Efron, B., and Tibshirani, R.,1986. Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. Statistical Science, 1(1), 54–75. http://www.jstor.org/stable/2245500

Einstein, H.H., Veneziano, D., Baecher, G.B., O'Reilly, K.J., 1983. The effect of discontinuity persistence on rock slope stability. International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts 20, 227–236. https://doi.org/10.1016/0148-9062(83)90003-7

Eivazy, H., Esmaieli, K., Jean, R., 2017. Modelling Geomechanical Heterogeneity of Rock Masses Using Direct and Indirect Geostatistical Conditional Simulation Methods. Rock Mechanics and Rock Engineering 50, 3175–3195. https://doi.org/10.1007/s00603-017-1293-0

Elith, J., and Leathwick, J. R., 2009. Species distribution models: ecological explanation and prediction across space and time. Annual review of ecology, evolution, and systematics, 40, 677-697.

Elith, J., Burgman, M.A., Regan, H.M., 2002. Mapping epistemic uncertainties and vague concepts in predictions of species distribution. Ecological Modelling 157, 313–329. https://doi.org/10.1016/S0304-3800(02)00202-8

Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann, F., R. Leathwick, J., Lehmann, A., Li, J., G. Lohmann, L., A. Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., McC. M. Overton, J., Townsend Peterson, A., J. Phillips, S., Richardson, K., Scachetti-Pereira, R., E. Schapire, R., Soberón, J., Williams, S., S. Wisz, M., E. Zimmermann, N., 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29, 129–151. https://doi.org/10.1111/j.2006.0906-7590.04596.x

Elith, J., Kearney, M., Phillips, S., 2010. The art of modelling range-shifting species. Methods in Ecology and Evolution 1, 330–342. https://doi.org/10.1111/j.2041-210X.2010.00036.x

Ellero, A., Loprieno, A., 2017. Nappe stack of Piemonte–Ligurian units south of Aosta Valley: New evidence from Urtier Valley (Western Alps). Geological Journal 53, 1665–1684. https://doi.org/10.1002/gj.2984

Engler, R., Guisan, A., Rechsteiner, L., 2004. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. Journal of Applied Ecology 41, 263–274. https://doi.org/10.1111/j.0021-8901.2004.00881.x

Ermini, L., Catani, F., Casagli, N., 2005. Artificial Neural Networks applied to landslide susceptibility assessment. Geomorphology 66, 327–343. https://doi.org/10.1016/j.geomorph.2004.09.025

Evans, J.P., Forster, C.B., Goddard, J.V., 1997. Permeability of fault-related rocks, and implications for hydraulic structure of fault zones. Journal of Structural Geology 19, 1393–1404. https://doi.org/10.1016/S0191-8141(97)00057-6

Fabbri, A.G., Chung, C.-J.F., Cendrero, A., Remondo, J., 2003. Is Prediction of Future Landslides Possible with a GIS? Natural Hazards 30, 487–503. https://doi.org/10.1023/B:NHAZ.0000007282.62071.75

Falaschi, F., Giacomelli, F., Federici, P.R., Puccinelli, A., D'Amato Avanzi, G., Pochini, A., Ribolini, A., 2009. Logistic regression versus artificial neural networks: landslide susceptibility evaluation in a sample area of the Serchio River valley, Italy. Natural Hazards 50, 551–569. https://doi.org/10.1007/s11069-009-9356-5

Fan, J. C., Yang, C. H., Chang, S. C., Huang, H. Y., & Guo, J. J. (2013). Effects of climate change on the potential of the landslides in the basin of Kaoping stream. J. Chin. Soil Water Conserv, 44(4), 335-350.

Fang, Z., Wang, Y., Peng, L., Hong, H., 2020. Integration of convolutional neural network and conventional machine learning classifiers for landslide susceptibility mapping. Computers & Geosciences 139, 104470. https://doi.org/10.1016/j.cageo.2020.104470

Fanos, A.M., Pradhan, B., 2019. A Novel Hybrid Machine Learning-Based Model for Rockfall Source Identification in Presence of Other Landslide Types Using LiDAR and GIS. Earth Systems and Environment 3, 491–506. https://doi.org/10.1007/s41748-019-00114-z

Feizizadeh, B., Blaschke, T., 2013. Land suitability analysis for Tabriz County, Iran: a multi-criteria evaluation approach using GIS. null 56, 1–23. https://doi.org/10.1080/09640568.2011.646964

Feizizadeh, B., Shadman Roodposhti, M., Jankowski, P., Blaschke, T., 2014. A GIS-based extended fuzzy multi-criteria evaluation for landslide susceptibility mapping. Computers & Geosciences 73, 208–221. https://doi.org/10.1016/j.cageo.2014.08.001

Ferrari, F., Apuani, T., Giani, G.P., 2012. Analisi spaziale e previsionale delle proprietà geomeccaniche degli ammassi rocciosi della Valle San Giacomo (SO), mediante tecniche geostatistiche, in: Geoitalia 2011. Pàtron Editore, pp. 21–30.

Ferrari, F., Apuani, T., Giani, G.P., 2014. Rock Mass Rating spatial estimation by geostatistical analysis. International Journal of Rock Mechanics and Mining Sciences 70, 162–176. https://doi.org/10.1016/j.ijrmms.2014.04.016

Ferrari, F., Ziegler, M., Apuani, T., Loew, S., 2019. Geostatistical analyses of exfoliation and tectonic joint set spacing in alpine granites (Aar Valley, Switzerland). Bulletin of Engineering Geology and the Environment 78, 1645–1668. https://doi.org/10.1007/s10064-018-1251-4

Ferretti, A., Fumagalli, A., Novali, F., Prati, C., Rocca, F., Rucci, A., 2011. A New Algorithm for Processing Interferometric Data-Stacks: SqueeSAR. IEEE Transactions on Geoscience and Remote Sensing 49, 3460–3470. https://doi.org/10.1109/TGRS.2011.2124465

Ferretti, A., Prati, C., Rocca, F., 2001. Permanent scatterers in SAR interferometry. IEEE Transactions on Geoscience and Remote Sensing 39, 8–20. https://doi.org/10.1109/36.898661

Figueiras, A., Roca, J., Cadarso-Suárez, C., 2005. A bootstrap method to avoid the effect of concurvity in generalised additive models in time series studies of air pollution. Journal of epidemiology and community health 59, 881–4. https://doi.org/10.1136/jech.2004.026740

Filippa, G., Cremonese, E., Galvagno, M., Isabellon, M., Bayle, A., Choler, P., Carlson, B.Z., Gabellani, S., Morra di Cella, U., Migliavacca, M., 2019. Climatic Drivers of Greening Trends in the Alps. Remote Sensing 11. https://doi.org/10.3390/rs11212527

Fischer, E.M., Beyerle, U., Knutti, R., 2013. Robust spatially aggregated projections of climate extremes. Nature Climate Change 3, 1033–1038. https://doi.org/10.1038/nclimate2051

Fotheringham, A. S., Brunsdon C., and Charlton M., 2002. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships, Wiley, Chichester.

Frattini, P., Crosta, G., Carrara, A., 2010. Techniques for evaluating the performance of landslide susceptibility models. Engineering Geology 111, 62–72. https://doi.org/10.1016/j.enggeo.2009.12.004

Frattini, P., Crosta, G., Carrara, A., Agliardi, F., 2008. Assessment of rockfall susceptibility by integrating statistical and physically-based approaches. Geomorphology 94, 419–437. https://doi.org/10.1016/j.geomorph.2006.10.037

Frattini, P., Crosta, G.B., Rossini, M., Allievi, J., 2018. Activity and kinematic behaviour of deep-seated landslides from PS-InSAR displacement rate measurements. Landslides 15, 1053–1070. https://doi.org/10.1007/s10346-017-0940-6

Frayssines, M., Hantz, D., 2006. Failure mechanisms and triggering factors in calcareous cliffs of the Subalpine Ranges (French Alps). Engineering Geology 86, 256–270. https://doi.org/10.1016/j.enggeo.2006.05.009

Frey M., Desmons J. And Neubauer F., 1999. The new metamorphic map of the Alps. Schweiz. mineral. petrogr. Mitt., v. 79, pp. 1-230.

Galli, A., Le Bayon, B., Schmidt, M.W., Burg, J.-P., Reusser, E., 2013. Tectonometamorphic history of the Gruf complex (Central Alps): exhumation of a granulite–migmatite complex with the Bergell pluton. Swiss Journal of Geosciences 106, 33–62. https://doi.org/10.1007/s00015-013-0120-1

Galli, M., Ardizzone, F., Cardinali, M., Guzzetti, F., Reichenbach, P., 2008. Comparing landslide inventory maps. Geomorphology 94, 268–289. https://doi.org/10.1016/j.geomorph.2006.09.023

Gariano, S.L., Guzzetti, F., 2016. Landslides in a changing climate. Earth-Science Reviews 162, 227–252. https://doi.org/10.1016/j.earscirev.2016.08.011

Gassner, C.; Promper, C.; Beguería, S.; Glade, T., 2015. Climate Change Impact for Spatial Landslide Susceptibility. In Engineering Geology for Society and Territory; Springer International Publishing: Cham, Switzerland; Vol. 1, 429–433.

Gigli, G., Morelli, S., Fornera, S., Casagli, N., 2014. Terrestrial laser scanner and geomechanical surveys for the rapid evaluation of rock fall susceptibility scenarios. Landslides 11, 1–14. https://doi.org/10.1007/s10346-012-0374-0

Giordan, D., Cignetti, M., Wrzesniak, A., Allasia, P., Bertolo, D., 2018. Operative Monographies: Development of a New Tool for the Effective Management of Landslide Risks. Geosciences 8. https://doi.org/10.3390/geosciences8120485

Gischig, V., Amann, F., Moore, J.R., Loew, S., Eisenbeiss, H., Stempfhuber, W., 2011. Composite rock slope kinematics at the current Randa instability, Switzerland, based on remote sensing and numerical modeling. Engineering Geology 118, 37–53. https://doi.org/10.1016/j.enggeo.2010.11.006

Giusti, F., Dal Piaz, G., Massironi, M., Schiavo, A., 2004. Carta geotettonica della Valle d'Aosta. Memorie di Scienze Geologiche 55 (2003), 129–149.

Glade, T., 2003. Landslide occurrence as a response to land use change: a review of evidence from New Zealand. CATENA 51, 297–314. https://doi.org/10.1016/S0341-8162(02)00170-4

Gobiet, A., Kotlarski, S., Beniston, M., Heinrich, G., Rajczak, J., Stoffel, M., 2014. 21st century climate change in the European Alps—A review. Science of The Total Environment 493, 1138–1151. https://doi.org/10.1016/j.scitotenv.2013.07.050

Goetz, J.N., Brenning, A., Petschko, H., Leopold, P., 2015. Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. Computers & geosciences 81, 1–11.

Goetz, J.N., Guthrie, R.H., Brenning, A., 2011. Integrating physical and empirical landslide susceptibility models using generalized additive models. Geomorphology 129, 376–386.

Goldstein, B.A., Polley, E.C., Briggs, F.B.S., 2011. Random forests for genetic association studies. Stat Appl Genet Mol Biol 10, 32–32. https://doi.org/10.2202/1544-6115.1691

Gómez, H., Kavzoglu, T., 2005. Assessment of shallow landslide susceptibility using artificial neural networks in Jabonosa River Basin, Venezuela. Engineering Geology 78, 11–27. https://doi.org/10.1016/j.enggeo.2004.10.004

Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, Oxford.

Grämiger, L.M., Moore, J.R., Gischig, V.S., Ivy-Ochs, S., Loew, S., 2017. Beyond debuttressing: Mechanics of paraglacial rock slope damage during repeat glacial cycles. Journal of Geophysical Research: Earth Surface 122, 1004–1036. https://doi.org/10.1002/2016JF003967

Green, D. M., & Swets, J. A., 1966. Signal detection theory and psychophysics (Vol. 1, pp. 1969-12). New York: Wiley.

Grömping, U., 2009. Variable Importance Assessment in Regression: Linear Regression versus Random Forest. null 63, 308–319. https://doi.org/10.1198/tast.2009.08199

Gunzburger, Y., Merrien-Soukatchoff, V., Guglielmi, Y., 2005. Influence of daily surface temperature fluctuations on rock slope stability: case study of the Rochers de Valabres slope (France). International Journal of Rock Mechanics and Mining Sciences, 42(3), 331-349. https://doi.org/10.1016/j.ijrmms.2004.11.003

Guo, J., Zhao, H., Ma, F., Li, K., Zhao, C., 2015. Investigating the permeability of fractured rock masses and the origin of water in a mine tunnel in Shandong Province, China. Water Science and Technology 72, 2006–2017. https://doi.org/10.2166/wst.2015.421

Guyennon, N., Valt, M., Salerno, F., Petrangeli, A.B., Romano, E., 2019. Estimating the snow water equivalent from snow depth measurements in the Italian Alps. Cold Regions Science and Technology 167, 102859. https://doi.org/10.1016/j.coldregions.2019.102859

Guzzetti, F., 2000. Landslide fatalities and the evaluation of landslide risk in Italy. Engineering Geology 58, 89–107. https://doi.org/10.1016/S0013-7952(00)00047-8

Guzzetti, F., Cardinali, M., Reichenbach, P., Cipolla, F., Sebastiani, C., Galli, M., Salvati, P., 2004. Landslides triggered by the 23 November 2000 rainfall event in the Imperia Province, Western Liguria, Italy. Engineering Geology 73, 229–245. https://doi.org/10.1016/j.enggeo.2004.01.006

Guzzetti, F., Carrara, A., Cardinali, M., Reichenbach, P., 1999. Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. Geomorphology 31, 181–216. https://doi.org/10.1016/S0169-555X(99)00078-1

Guzzetti, F., Mondini, A.C., Cardinali, M., Fiorucci, F., Santangelo, M., Chang, K.-T., 2012. Landslide inventory maps: New tools for an old problem. Earth-Science Reviews 112, 42–66. https://doi.org/10.1016/j.earscirev.2012.02.001

Guzzetti, F., Peruccacci, S., Rossi, M., Stark, C.P., 2007. Rainfall thresholds for the initiation of landslides in central and southern Europe. Meteorology and Atmospheric Physics 98, 239–267. https://doi.org/10.1007/s00703-007-0262-7

Guzzetti, F., Peruccacci, S., Rossi, M., Stark, C.P., 2008. The rainfall intensity–duration control of shallow landslides and debris flows: an update. Landslides 5, 3–17. https://doi.org/10.1007/s10346-007-0112-1

Guzzetti, F., Reichenbach, P., Ardizzone, F., Cardinali, M., Galli, M., 2006. Estimating the quality of landslide susceptibility models. Geomorphology 81, 166–184. https://doi.org/10.1016/j.geomorph.2006.04.007

Guzzetti, F., Reichenbach, P., Cardinali, M., Galli, M., Ardizzone, F., 2005. Probabilistic landslide hazard assessment at the basin scale. Geomorphology 72, 272–299. https://doi.org/10.1016/j.geomorph.2005.06.002

Hales, T.C., Roering, J.J., 2007. Climatic controls on frost cracking and implications for the evolution of bedrock landscapes. Journal of Geophysical Research: Earth Surface 112. https://doi.org/10.1029/2006JF000616

Hall, K., Thorn, C., Sumner, P., 2012. On the persistence of 'weathering.' Geomorphology 149–150, 1–10. https://doi.org/10.1016/j.geomorph.2011.12.024

Hantz, D., Vengeon, J.M., Dussauge-Peisser, C., 2003. An historical, geomechanical and probabilistic approach to rock-fall hazard assessment. Natural Hazards and Earth System Sciences 3, 693–701. https://doi.org/10.5194/nhess-3-693-2003

Hartkamp, A.D., De Beurs, K., Stein, A. and White, J.W., 1999. Interpolation Techniques for Climate Variables, CIMMYT, Mexico, D.F.

Hartmeyer, I., Delleske, R., Keuschnig, M., Krautblatter, M., Lang, A., Schrott, L., Otto, J.-C., 2020. Current glacier recession causes significant rockfall increase: the immediate paraglacial response of deglaciating cirque walls. Earth Surface Dynamics 8, 729–751. https://doi.org/10.5194/esurf-8-729-2020

Hastie, T.J., & Tibshirani, R.J. (1990). Generalized Additive Models (1st ed.). Routledge. https://doi.org/10.1201/9780203753781

He, K., Li, Y., Ma, G., Hu, X., Liu, B., Ma, Z., Xu, Z., 2021. Failure mode analysis of post-seismic rockfall in shattered mountains exemplified by detailed investigation and numerical modelling. Landslides 18, 425–446. https://doi.org/10.1007/s10346-020-01532-1

He, Z.H., Parajka, J., Tian, F.Q., Blöschl, G., 2014. Estimating degree-day factors from MODIS for snowmelt runoff modeling. Hydrology and Earth System Sciences 18, 4773–4789. https://doi.org/10.5194/hess-18-4773-2014

Heleno, S., Matias, M., Pina, P., Sousa, A.J., 2016. Semiautomated object-based classification of rain-induced landslides with VHR multispectral images on Madeira Island. Natural Hazards and Earth System Sciences 16, 1035–1048. https://doi.org/10.5194/nhess-16-1035-2016

Herrera, G., Gutiérrez, F., García-Davalillo, J.C., Guerrero, J., Notti, D., Galve, J.P., Fernández-Merodo, J.A., Cooksley, G., 2013. Multi-sensor advanced DInSAR monitoring of very slow landslides: The Tena Valley case study (Central Spanish Pyrenees). Remote Sensing of Environment 128, 31–43. https://doi.org/10.1016/j.rse.2012.09.020

Hijmans, R.J., 2013. Dismo, species distribution modelling. R package version 1.0-15. Available online: https://cran.r-project.org/web/packages/dismo/dismo.pdf

Hoek E., 2000. Analysis of rockfall hazards. In: Practical rock engineering, pp 117–136

Hofstra, N., Haylock, M., New, M., Jones, P., Frei, C., 2008. Comparison of six methods for the interpolation of daily, European climate data. Journal of Geophysical Research: Atmospheres 113. https://doi.org/10.1029/2008JD010100

Hölbling, D., Eisank, C., Albrecht, F., Vecchiotti, F., Friedl, B., Weinke, E., Kociu, A., 2017. Comparing Manual and Semi-Automated Landslide Mapping Based on Optical Satellite Images from Different Sensors. Geosciences 7. https://doi.org/10.3390/geosciences7020037

Hölbling, D., Friedl, B., Eisank, C., 2015. An object-based approach for semi-automated landslide change detection and attribution of changes to landslide classes in northern Taiwan. Earth Science Informatics 8, 327–335. https://doi.org/10.1007/s12145-015-0217-3

Hong, H., Miao, Y., Liu, J., Zhu, A.-X., 2019. Exploring the effects of the design and quantity of absence data on the performance of random forest-based landslide susceptibility mapping. CATENA 176, 45–64. https://doi.org/10.1016/j.catena.2018.12.035

Hong, H., Naghibi, S.A., Pourghasemi, H.R., Pradhan, B., 2016. GIS-based landslide spatial modeling in Ganzhou City, China. Arabian Journal of Geosciences 9, 112. https://doi.org/10.1007/s12517-015-2094-y

Hosmer, D. W. and Lemeshow, S., 2000. Applied logistic regression, Wiley, New York.

Hosmer, D.W., Lemeshow, S., Sturdivant, R.X., 2013. Applied Logistic Regression, 3rd ed, Wiley Series in Probability and Statistics. Wiley. https://doi.org/10.1002/9781118548387

Huabin, W., Gangjun, L., Weiya, X., Gonghui, W., 2005. GIS-based landslide hazard assessment: an overview. Progress in Physical Geography: Earth and Environment 29, 548–567. https://doi.org/10.1191/0309133305pp462ra

Huang, F., Ye, Z., Jiang, S.-H., Huang, J., Chang, Z., Chen, J., 2021. Uncertainty study of landslide susceptibility prediction considering the different attribute interval numbers of environmental factors and different data-based models. CATENA 202, 105250. https://doi.org/10.1016/j.catena.2021.105250

Huber RK, Marquer D., 1998. The tectonometamorphic history of the peridotitic Chiavenna unit from Mesozoic to Tertiary tectonics:a restoration controlled by melt polarity indicators (Eastern Swiss Alps). Tectonophysics 296:205–223

Hudson JA, Harrison JP., 1997. Anisotropy and inhomogeneity. In Engineering Rock Mechanics. Pergamon: Oxford; 163–172.

Huisman, M., Hack, H.R.G.K., Nieuwenhuis, J.D., 2006. Predicting Rock Mass Decay in Engineering Lifetimes: The Influence of Slope Aspect and Climate. Environmental and Engineering Geoscience 12, 39–51. https://doi.org/10.2113/12.1.39

Hussain, M., Chen, D., Cheng, A., Wei, H., Stanley, D., 2013. Change detection from remotely sensed images: From pixel-based to object-based approaches. ISPRS Journal of Photogrammetry and Remote Sensing 80, 91–106. https://doi.org/10.1016/j.isprsjprs.2013.03.006

Hussin, H.Y., Zumpano, V., Reichenbach, P., Sterlacchini, S., Micu, M., van Westen, C., Bălteanu, D., 2016. Different landslide sampling strategies in a grid-based bi-variate statistical susceptibility model. Geomorphology 253, 508–523. https://doi.org/10.1016/j.geomorph.2015.10.030

Hutchinson, M.F., 1995. Interpolating mean rainfall using thin plate smoothing splines. International Journal of Geographical Information Systems, 9(4): 385-403.

Hutter, F., Hoos, H.H., Leyton-Brown, K., 2011. Sequential Model-Based Optimization for General Algorithm Configuration, in: Coello, C.A.C. (Ed.), Learning and Intelligent Optimization. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 507–523.

IPCC, 2007. Climate change 2007: the physical science basis. In: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., Miller, H.L. (Eds.), Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

IPCC, 2013. Climate change 2013: the physical science basis. In: Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Midgley, P.M. (Eds.), Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Isaaks, E.H. and Srivastava, R.M., 1989. Applied Geostatistics. Oxford University Press, New York, 561 pp.

Ishikawa, M., Kurashige, Y., Hirakawa, K., 2004. Analysis of crack movements observed in an alpine bedrock cliff. Earth Surface Processes and Landforms 29, 883–891. https://doi.org/10.1002/esp.1076

ISPRA Ambiente IFFI Catalogue. Available online: http://www.isprambiente.gov.it/it/progetti/suolo-eterritorio-1/iffi-inventario-dei-fenomeni-franosi-in-italia

ISRM, 1978. Suggested methods for the quantitative description of discontinuities in rock masses. Commission on the standardization of Laboratory and Field Tests in Rock Mechanics, ISRM.

Jaboyedoff, M., Baillifard, F., Bardou, E., Girod, F., 2004. The effect of weathering on Alpine rock instability. Quarterly Journal of Engineering Geology and Hydrogeology 37, 95–103. https://doi.org/10.1144/1470-9236/03-046

Jaboyedoff, M., Crosta, G.B., Stead, D., 2011. Slope tectonics: a short introduction. Geological Society, London, Special Publications 351, 1–10. https://doi.org/10.1144/SP351.1

Jacobs, L., Dewitte, O., Poesen, J., Maes, J., Mertens, K., Sekajugo, J., Kervyn, M., 2017. Landslide characteristics and spatial distribution in the Rwenzori Mountains, Uganda. Journal of African Earth Sciences 134, 917–930. https://doi.org/10.1016/j.jafrearsci.2016.05.013

Jacobs, L., Kervyn, M., Reichenbach, P., Rossi, M., Marchesini, I., Alvioli, M., Dewitte, O., 2020. Regional susceptibility assessments with heterogeneous landslide information: Slope unit- vs. pixel-based approach. Geomorphology 356, 107084. https://doi.org/10.1016/j.geomorph.2020.107084

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. Statistical Learning, in: An Introduction to Statistical Learning, Springer Texts in Statistics. Springer New York, New York, NY, pp. 15–57. https://doi.org/10.1007/978-1-4614-7138-7_2

Janitza, S., Binder, H., & Boulesteix, A. L., 2016. Pitfalls of hypothesis tests and model selection on bootstrap samples: Causes and consequences in biometrical applications. Biometrical journal. 58(3), 447–473. https://doi.org/10.1002/bimj.201400246

Javadinejad, S., Dara, R., Jafary, F., 2020. Climate change scenarios and effects on snowmelt runoff. Civ. Eng. J. 6, 1715–1725. https://doi.org/10.28991/cej-2020-03091577.

Jia, G., Tian, Y., Liu, Y., Zhang, Y., 2008. A static and dynamic factors-coupled forecasting model of regional rainfall-induced landslides: A case study of Shenzhen. Science in China Series E: Technological Sciences 51, 164–175. https://doi.org/10.1007/s11431-008-6013-2

Jollifee, I. T. and Stephenson, D. B., 2003. Forecast verification, A Practitioner's Guide in Atmospheric Science, John Wiley, Chichester, UK, 240 pp.

Journel, A.G. and Huijbregts, C.J., 1978. Mining Geostatistics. Academic Press, London, 600 pp.

Juang C.S., Stanley T.A., Kirschbaum D.B., 2019. Using citizen science to expand the global map of landslides: Introducing the Cooperative Open Online Landslide Repository (COOLR). PLoS ONE 14(7): e0218657. https://doi.org/10.1371/journal.pone.0218657

Kalantar, B., Pradhan, B., Naghibi, S.A., Motevalli, A., Mansor, S., 2018. Assessment of the effects of training data selection on the landslide susceptibility mapping: a comparison between support vector machine (SVM), logistic regression (LR) and artificial neural networks (ANN). null 9, 49–69. https://doi.org/10.1080/19475705.2017.1407368

Keefer, D.K., 1984a. Landslides caused by earthquakes. Geol Soc America Bull 95, 406. https://doi.org/10.1130/0016-7606(1984)95<406:LCBE>2.0.CO;2

Keefer, D.K., 1984b. Rock Avalanches Caused by Earthquakes: Source Characteristics. Science 223, 1288–1290. https://doi.org/10.1126/science.223.4642.1288

Kegang, L., Lin, M., Xiangxing, L., Shoujian, P., 2016. Effect of Drying-Wetting Cycles on Triaxial Compression Mechanical Properties of Sandstone. Journal of Engineering Science & Technology Review 9.

Kim, H.G., Lee, D.K., Park, C., Ahn, Y., Kil, S.-H., Sung, S., Biging, G.S., 2018. Estimating landslide susceptibility areas considering the uncertainty inherent in modeling methods. Stochastic Environmental Research and Risk Assessment 32, 2987–3019. https://doi.org/10.1007/s00477-018-1609-y

Kim, H.G., Lee, D.K., Park, C., Kil, S., Son, Y., Park, J.H., 2015. Evaluating landslide hazards using RCP 4.5 and 8.5 scenarios. Environ. Earth Sci. 73, 1385–1400. https://doi.org/10.1007/s12665-014-3775-7.

Kiraly, L., 1969. Anisotropie et hétérogénéité de la perméabilité dans les calcaires fissurés (Anisotropy and heterogeneity of permeability in fractured limestones). Eclogae Geol. Helv. 62/2, 613–619.

Kirchner, M., Faus-Kessler, T., Jakobi, G., Leuchner, M., Ries, L., Scheel, H.-E., Suppan, P., 2013. Altitudinal temperature lapse rates in an Alpine valley: trends and the influence of season and weather patterns. International Journal of Climatology 33, 539–555. https://doi.org/10.1002/joc.3444

Knevels, R., Petschko, H., Proske, H., Leopold, P., Maraun, D., Brenning, A., 2020. Event-Based Landslide Modeling in the Styrian Basin, Austria: Accounting for Time-Varying Rainfall and Land Cover. Geosciences 10. https://doi.org/10.3390/geosciences10060217

Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence, vol. 2, pp. 1137-1143.

Komisarczyk K., Kozminski P., Maksymiuk S. and Biecek P., 2021. treeshap: Fast SHAP values computation for ensemble models. R package version 0.0.1. https://github.com/ModelOriented/treeshap

Kouli, M., Loupasakis, C., Soupios, P., Rozos, D., Vallianatos, F., 2014. Landslide susceptibility mapping by comparing the WLC and WofE multi-criteria methods in the West Crete Island, Greece. Environmental Earth Sciences 72, 5197–5219. https://doi.org/10.1007/s12665-014-3389-0

Kovacs, G., 1981. Seepage hydraulics. Seepage hydraulics. Elsivier Publishing Company.

Krautblatter, M., Moore, J.R., 2014. Rock slope instability and erosion: toward improved process understanding. Earth Surface Processes and Landforms 39, 1273–1278. https://doi.org/10.1002/esp.3578

Krautblatter, M., Moser, M., 2009. A nonlinear model coupling rockfall and rainfall intensity based \newline on a four year measurement in a high Alpine rock wall (Reintal, German Alps). Natural Hazards and Earth System Sciences 9, 1425–1432. https://doi.org/10.5194/nhess-9-1425-2009

Krige, D.G., 1951. A Statistical Approaches to Some Basic Mine Valuation Problems on the Witwatersrand. Journal of the Chemical, Metallurgical and Mining Society of South Africa, 52, 119-139.

Laceby, J.P., Chartin, C., Evrard, O., Onda, Y., Garcia-Sanchez, L., Cerdan, O., 2016. Rainfall erosivity in catchments contaminated with fallout from the Fukushima Daiichi nuclear power plant accident. Hydrology and Earth System Sciences 20, 2467–2482. https://doi.org/10.5194/hess-20-2467-2016

Lagomarsino, D., Tofani, V., Segoni, S., Catani, F., Casagli, N., 2017. A Tool for Classification and Regression Using Random Forest Methodology: Applications to Landslide Susceptibility Mapping and Soil Thickness Modeling. Environmental Modeling & Assessment 22, 201–214. https://doi.org/10.1007/s10666-016-9538-y

Leclère, H., Cappa, F., Faulkner, D., Fabbri, O., Armitage, P., Blake, O., 2015. Development and maintenance of fluid overpressures in crustal fault zones by elastic compaction and implications for earthquake swarms. Journal of Geophysical Research: Solid Earth 120, 4450–4473. https://doi.org/10.1002/2014JB011759

Leith, K., Moore, J.R., Amann, F., Loew, S., 2014. In situ stress control on microcrack generation and macroscopic extensional fracture in exhuming bedrock. J. Geophys. Res. Solid Earth 119, 594–615. https://doi.org/10.1002/2012JB009801

Leith, K.J., 2012. Stress development and geomechanical controls on the geomorphic evolution of alpine valleys. ETH Zurich.

Leonarduzzi, E., Molnar, P., 2020. Data limitations and potential of hourly and daily rainfall thresholds for shallow landslides. Natural Hazards and Earth System Sciences Discussions 2020, 1–25. https://doi.org/10.5194/nhess-2020-125

Li, J. and Heap, A.D., 2008. A Review of Spatial Interpolation Methods for Environmental Scientists. Geoscience Australia, Record 2008/23, 137 pp.

Li, J., Heap, A.D., 2011. A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. Ecological Informatics 6, 228–241. https://doi.org/10.1016/j.ecoinf.2010.12.003

Li, J., Heap, A.D., 2014. Spatial interpolation methods applied in the environmental sciences: A review. Environmental Modelling & Software 53, 173–189. https://doi.org/10.1016/j.envsoft.2013.12.008

Li, J., Tian, L., Wang, Y., Jin, S., Li, T., Hou, X., 2021a. Optimal sampling strategy of water quality monitoring at high dynamic lakes: A remote sensing and spatial simulated annealing integrated

approach. Science of The Total Environment 777, 146113. https://doi.org/10.1016/j.scitotenv.2021.146113

Li, J., Tian, L., Wang, Y., Jin, S., Li, T., Hou, X., 2021b. Optimal sampling strategy of water quality monitoring at high dynamic lakes: A remote sensing and spatial simulated annealing integrated approach. Science of The Total Environment 777, 146113. https://doi.org/10.1016/j.scitotenv.2021.146113

Li, R., Shinde, A., Liu, A., Glaser, S., Lyou, Y., Yuh, B., Wong, J., Amini, A., 2020. Machine Learning–Based Interpretation and Visualization of Nonlinear Interactions in Prostate Cancer Survival. JCO Clinical Cancer Informatics 637–646. https://doi.org/10.1200/CCI.20.00002

Liati, A., Gebauer, D., Fanning, C.M., 2003. The youngest basic oceanic magmatism in the Alps (Late Cretaceous ; Chiavenna unit, Central Alps): geochronological constraints and geodynamic significance. Contributions to Mineralogy and Petrology 146, 144–158. https://doi.org/10.1007/s00410-003-0485-7

Lin, G.-F., Chang, M.-J., Huang, Y.-C., Ho, J.-Y., 2017. Assessment of susceptibility to rainfall-induced landslides using improved self-organizing linear output map, support vector machine, and logistic regression. Engineering Geology 224, 62–74. https://doi.org/10.1016/j.enggeo.2017.05.009

Lin, Y., Jeon, Y., 2006. Random Forests and Adaptive Nearest Neighbors. null 101, 578–590. https://doi.org/10.1198/016214505000001230

Liu Y., Just A., 2021. SHAPforxgboost: SHAP Plots for 'XGBoost'. R package version 0.1.1. https://CRAN.R-project.org/package=SHAPforxgboost

Lombardo, L., Bakka, H., Tanyas, H., van Westen, C., Mai, P.M., Huser, R., 2019a. Geostatistical Modeling to Capture Seismic-Shaking Patterns From Earthquake-Induced Landslides. Journal of Geophysical Research: Earth Surface 124, 1958–1980. https://doi.org/10.1029/2019JF005056

Lombardo, L., Cama, M., Maerker, M., Rotigliano, E., 2014. A test of transferability for landslides susceptibility models under extreme climatic events: application to the Messina 2009 disaster. Natural Hazards 74, 1951–1989. https://doi.org/10.1007/s11069-014-1285-2

Lombardo, L., Mai, P.M., 2018. Presenting logistic regression-based landslide susceptibility results. Engineering Geology 244, 14–24. https://doi.org/10.1016/j.enggeo.2018.07.019

Lombardo, L., Opitz, T., Ardizzone, F., Guzzetti, F., Huser, R., 2020. Space-time landslide predictive modelling. Earth-Science Reviews 209, 103318. https://doi.org/10.1016/j.earscirev.2020.103318

Lombardo, L., Opitz, T., Huser, R., 2018. Point process-based modeling of multiple debris flow landslides using INLA: an application to the 2009 Messina disaster. Stochastic Environmental Research and Risk Assessment 32, 2179–2198. https://doi.org/10.1007/s00477-018-1518-0

Lombardo, L., Opitz, T., Huser, R., 2019b. 3 - Numerical Recipes for Landslide Spatial Prediction Using R-INLA: A Step-by-Step Tutorial, in: Pourghasemi, H.R., Gokceoglu, C. (Eds.), Spatial Modeling in GIS and R for Earth and Environmental Sciences. Elsevier, pp. 55–83. https://doi.org/10.1016/B978-0-12-815226-3.00003-X

Lorentz, J.F., Calijuri, M.L., Marques, E.G., Baptista, A.C., 2016. Multicriteria analysis applied to landslide susceptibility mapping. Natural Hazards 83, 41–52. https://doi.org/10.1007/s11069-016-2300-6

Losasso, L., Sdao, F., 2018. The artificial neural network for the rockfall susceptibility assessment. A case study in Basilicata (Southern Italy). Geomatics, Natural Hazards and Risk 9, 737–759. https://doi.org/10.1080/19475705.2018.1476413

Loye, A., Pedrazzini, A., Theule, J.I., Jaboyedoff, M., Liébault, F., Metzger, R., 2012. Influence of bedrock structures on the spatial pattern of erosional landforms in small alpine catchments: STRUCTURAL CONTROL OF EROSION. Earth Surf. Process. Landforms 37, 1407–1423. https://doi.org/10.1002/esp.3285

Lu, P., Stumpf, A., Kerle, N., Casagli, N., 2011. Object-Oriented Change Detection for Landslide Rapid Mapping. IEEE Geoscience and Remote Sensing Letters 8, 701–705. https://doi.org/10.1109/LGRS.2010.2101045

Lubo-Robles, D., Devegowda, D., Jayaram, V., Bedle, H., Marfurt, K.J., Pranter, M.J., 2020. Machine learning model interpretability using SHAP values: Application to a seismic facies classification task. Presented at the SEG International Exposition and Annual Meeting. https://doi.org/10.1190/segam2020-3428275.1

Lucas, D., Herzog, R., Iten, M., Buschor, H., Kieper, A., Askarinejad, A., Springman, S.M., 2020. Modelling of landslides in a scree slope induced by groundwater and rainfall. International Journal of Physical Modelling in Geotechnics 20, 177–197. https://doi.org/10.1680/jphmg.18.00106

Luino, F., De Graff, J., Roccati, A., Biddoccu, M., Cirio, C.G., Faccini, F., Turconi, L., 2020. Eighty Years of Data Collected for the Determination of Rainfall Threshold Triggering Shallow Landslides and Mud-Debris Flows in the Alps. Water 12. https://doi.org/10.3390/w12010133

Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.-I., 2019. Explainable AI for Trees: From Local Explanations to Global Understanding. arXiv:1905.04610.

Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.-I., 2020. From local explanations to global understanding with explainable AI for trees. Nat Mach Intell 2, 56–67. https://doi.org/10.1038/s42256-019-0138-9

Lundberg, S.M., Lee, S.-I., 2017. A unified approach to interpreting model predictions, in: Proceedings of the 31st International Conference on Neural Information Processing Systems. pp. 4768–4777.

Ma, T., Brus, D.J., Zhu, A.-X., Zhang, L., Scholten, T., 2020. Comparison of conditioned Latin hypercube and feature space coverage sampling for predicting soil classes using simulation from soil maps. Geoderma 370, 114366. https://doi.org/10.1016/j.geoderma.2020.114366

Macciotta, R., Hendry, M., Cruden, D.M., Blais-Stevens, A., Edwards, T., 2017. Quantifying rock fall probabilities and their temporal distribution associated with weather seasonality. Landslides 14, 2025–2039. https://doi.org/10.1007/s10346-017-0834-7

Macciotta, R., Martin, C.D., Edwards, T., Cruden, D.M., Keegan, T., 2015. Quantifying weather conditions for rock fall hazard management. null 9, 171–186. https://doi.org/10.1080/17499518.2015.1061673

Magnin, F., Westermann, S., Pogliotti, P., Ravanel, L., Deline, P., Malet, E., 2017. Snow control on active layer thickness in steep alpine rock walls (Aiguille du Midi, 3842ma.s.l., Mont Blanc massif). CATENA 149, 648–662. https://doi.org/10.1016/j.catena.2016.06.006

Malamud, B.D., Turcotte, D.L., Guzzetti, F., Reichenbach, P., 2004. Landslide inventories and their statistical properties. Earth Surface Processes and Landforms 29, 687–711. https://doi.org/10.1002/esp.1064

Mangalathu, S., Hwang, S.-H., Jeon, J.-S., 2020. Failure mode and effects analysis of RC members based on machine-learning-based SHapley Additive exPlanations (SHAP) approach. Engineering Structures 219, 110927. https://doi.org/10.1016/j.engstruct.2020.110927

Marquer D., Baudin T., Peucat J-J, Persoz F., 1994. Rb-Sr mica ages in the Alpine shear zones of the Truzzo granite: timing of the Tertiary alpine P-T-deformations in the Tambò nappe (Central Alps, Switzerland). Eclogae Geol Helv 87(1):225–239

Martha, T.R., Kerle, N., van Westen, C.J., Jetten, V., Kumar, K.V., 2011. Segment Optimization and Data-Driven Thresholding for Knowledge-Based Landslide Detection by Object-Based Image Analysis. IEEE Transactions on Geoscience and Remote Sensing 49, 4928–4943. https://doi.org/10.1109/TGRS.2011.2151866

Martha, T.R., Kerle, N., Westen, C.J. van, Jetten, V., Kumar, K.V., 2012. Object-oriented analysis of multi-temporal panchromatic images for creation of historical landslide inventories. ISPRS Journal of Photogrammetry and Remote Sensing 67, 105–119. https://doi.org/10.1016/j.isprsjprs.2011.11.004

Martinello, C., Cappadonia, C., Conoscenti, C., Agnesi, V., Rotigliano, E., 2021. Optimal slope units partitioning in landslide susceptibility mapping. Journal of Maps 17, 152–162. https://doi.org/10.1080/17445647.2020.1805807

Martínez-Muñoz, G., Suárez, A., 2010. Out-of-bag estimation of the optimal sample size in bagging. Pattern Recognition 43, 143–152. https://doi.org/10.1016/j.patcog.2009.05.010

Marzorati, S., Luzi, L., De Amicis, M., 2002. Rock falls induced by earthquakes: a statistical approach. Soil Dynamics and Earthquake Engineering 22, 565–577. https://doi.org/10.1016/S0267-7261(02)00036-2

Matasci, B., Carrea, D., Jaboyedoff, M., Pedrazzini, A., Stock, G.M., Oppikofer, T., 2011. Structural characterization of rockfall sources in Yosemite Valley from remote sensing data: Toward more accurate susceptibility assessment, in: Proceedings of the 14th ISSMGE Pan-American Conference.

Matasci, B., Jaboyedoff, M., Loye, A., Pedrazzini, A., Derron, M.-H., Pedrozzi, G., 2015. Impacts of fracturing patterns on the rockfall susceptibility and erosion rate of stratified limestone. Geomorphology 241, 83–97. https://doi.org/10.1016/j.geomorph.2015.03.037

Matasci, B., Stock, G.M., Jaboyedoff, M., Carrea, D., Collins, B.D., Guérin, A., Matasci, G., Ravanel, L., 2018. Assessing rockfall susceptibility in steep and overhanging slopes using three-dimensional analysis of failure mechanisms. Landslides 15, 859–878. https://doi.org/10.1007/s10346-017-0911-y

Matheron, G., 1963. Principles of geostatistics, Econ. Geol., 58, 1246–1266.

Matsuoka, N., 2008. Frost weathering and rockwall erosion in the southeastern Swiss Alps: Long-term (1994–2006) observations. Geomorphology 99, 353–368. https://doi.org/10.1016/j.geomorph.2007.11.013

Matsuoka, N., 2013. Time-lapse photography applied to monitoring of alpine slope processes. Tsukuba geoenvironmental sciences 9, 13–19.

Matsuoka, N., 2019. A multi-method monitoring of timing, magnitude and origin of rockfall activity in the Japanese Alps. Geomorphology 336, 65–76. https://doi.org/10.1016/j.geomorph.2019.03.023

Mazzoccola D., 1993. La dinamica dei versanti della media Valchiavenna (SO): analisi geomeccanica dei fenomeni di stabilità in atto e potenziali. Università degli Studi di Milano. Unpublished Phd Thesis.

McColl, S.T., 2012. Paraglacial rock-slope stability. Geomorphology 153–154, 1–16. https://doi.org/10.1016/j.geomorph.2012.02.015

Meinhardt, M., Fink, M., Tünschel, H., 2015. Landslide susceptibility analysis in central Vietnam based on an incomplete landslide inventory: Comparison of a new method to calculate weighting factors by means of bivariate statistics. Geomorphology 234, 80–97. https://doi.org/10.1016/j.geomorph.2014.12.042

Melchiorre, C., Castellanos Abella, E.A., van Westen, C.J., Matteucci, M., 2011. Evaluation of prediction capability, robustness, and sensitivity in non-linear landslide susceptibility models, Guantánamo, Cuba. Computers & Geosciences 37, 410–425. https://doi.org/10.1016/j.cageo.2010.10.004

Merghadi, A., Yunus, A.P., Dou, J., Whiteley, J., ThaiPham, B., Bui, D.T., Avtar, R., Abderrahmane, B., 2020. Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. Earth-Science Reviews 207, 103225. https://doi.org/10.1016/j.earscirev.2020.103225

Mergili, M., Jaboyedoff, M., Pullarello, J., Pudasaini, S.P., 2020. Back calculation of the 2017 Piz Cengalo–Bondo landslide cascade with r.avaflow: what we can do and what we can learn. Natural Hazards and Earth System Sciences 20, 505–520. https://doi.org/10.5194/nhess-20-505-2020

Messenzehl, K., Meyer, H., Otto, J.-C., Hoffmann, T., Dikau, R., 2017. Regional-scale controls on the spatial activity of rockfalls (Turtmann Valley, Swiss Alps) — A multivariate modeling approach. Geomorphology 287, 29–45. https://doi.org/10.1016/j.geomorph.2016.01.008

Messenzehl, K., Viles, H., Otto, J.-C., Ewald, A., Dikau, R., 2018. Linking rock weathering, rockwall instability and rockfall supply on talus slopes in glaciated hanging valleys (Swiss Alps). Permafrost and Periglacial Processes 29, 135–151. https://doi.org/10.1002/ppp.1976

Meusburger, K., Alewell, C., 2008. Impacts of anthropogenic and environmental factors on the occurrence of shallow landslides in an alpine catchment (Urseren Valley, Switzerland). Natural Hazards and Earth System Sciences 8, 509–520. https://doi.org/10.5194/nhess-8-509-2008

Micheletti, N., Foresti, L., Robert, S., Leuenberger, M., Pedrazzini, A., Jaboyedoff, M., Kanevski, M., 2014. Machine Learning Feature Selection Methods for Landslide Susceptibility Mapping. Math Geosci 46, 33–57. https://doi.org/10.1007/s11004-013-9511-0

Michie, D., Spiegelhalter, D.J., Taylor, C.C., 1994. Machine learning, neural and statistical classification.

Michoud, C., Derron, M.-H., Horton, P., Jaboyedoff, M., Baillifard, F.-J., Loye, A., Nicolet, P., Pedrazzini, A., Queyrel, A., 2012. Rockfall hazard and risk assessments along roads at a regional scale: example in Swiss Alps. Natural Hazards and Earth System Sciences 12, 615–629. https://doi.org/10.5194/nhess-12-615-2012

Milsch, H., Priegnitz, M., Blöcher, G., 2011. Permeability of gypsum samples dehydrated in air. Geophysical Research Letters 38. https://doi.org/10.1029/2011GL048797

Minasny, B., McBratney, A.B., 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. Computers & Geosciences 32, 1378–1388. https://doi.org/10.1016/j.cageo.2005.12.009

Miščević, P., Vlastelica, G., 2014. Impact of weathering on slope stability in soft rock mass. Journal of Rock Mechanics and Geotechnical Engineering 6, 240–250. https://doi.org/10.1016/j.jrmge.2014.03.006

Mokhtari, K.E., Higdon, B.P., Başar, A., 2019. Interpreting Financial Time Series with SHAP Values, in: Proceedings of the 29th Annual International Conference on Computer Science and Software Engineering, CASCON '19. IBM Corp., USA, pp. 166–172.

Molnar, C. 2019. Interpretable machine learning. A Guide for Making Black Box Models Explainable. https://christophm.github.io/interpretable-ml-book/.

Mondini, A.C., Guzzetti, F., Reichenbach, P., Rossi, M., Cardinali, M., Ardizzone, F., 2011. Semi-automatic recognition and mapping of rainfall induced shallow landslides using optical satellite images. Remote Sensing of Environment 115, 1743–1757. https://doi.org/10.1016/j.rse.2011.03.006

Montrasio A., Sciesa E., 1988. Carta geologica della valle Spluga e aree adiacenti. Profilo Crosta Profonda 88 02, CNR, Dipartimento Scienze della Terra, Università degli Studi di Milano

Moore, J.R., Sanders, J.W., Dietrich, W.E., Glaser, S.D., 2009. Influence of rock mass strength on the erosion rate of alpine cliffs. Earth Surf. Process. Landforms 34, 1339–1352. https://doi.org/10.1002/esp.1821

Morcioni A., Apuani T., Cecinato F. (2022) – "The role of temperature in the stress-strain evolution of Alpine rock-slopes: thermo-mechanical modelling of the Cimaganda Rockslide". Accepted Paper (9/01/2022) RMRE-D-21-00708R1, in press on Rock Mechanics and Rock Engineering, Springer

Morcioni A., Bajni G., Apuani T., 2020. The Cimaganda Rockslide (Italian Alps): geomechanical characterization and hydro-mechanical numerical modelling. RENDICONTI ONLINE DELLA SOCIETÀ GEOLOGICA ITALIANA Vol. 52/2020, 40-46 https://doi.org/10.3301/ROL.2020.15

Morrow, C., Lockner, D., Hickman, S., Rusanov, M., Röckel, T., 1994. Effects of lithology and depth on the permeability of core samples from the Kola and KTB drill holes. Journal of Geophysical Research: Solid Earth 99, 7263–7274. https://doi.org/10.1029/93JB03458

Muenchow, J., Brenning, A., Richter, M., 2012. Geomorphic process rates of landslides along a humidity gradient in the tropical Andes. Geomorphology 139–140, 271–284. https://doi.org/10.1016/j.geomorph.2011.10.029

Murillo, M.L., Hunter, G.J., 1997. Assessing uncertainty due to elevation error in a landslide susceptibility model. Transactions in GIS 2, 289–298. https://doi.org/10.1111/j.1467-9671.1997.tb00058.x

Nahayo, L., Mupenzi, C., Habiyaremye, G., Kalisa, E., Udahogora, M., Nzabarinda, V., Li, L., 2019. Landslides Hazard Mapping in Rwanda Using Bivariate Statistical Index Method. Environmental Engineering Science 36, 892–902. https://doi.org/10.1089/ees.2018.0493

Najser, J., Sosna, K., Novakova, L., Broz, M., Kasikova, J., Michalkova, J., Novák, P., Vanecek, M., Zaruba, J., 2011. The relationship between matrix hydraulic conductivity and P-wave and S-wave ultrasound velocities in granites.

Neuhäuser, B., Damm, B., Terhorst, B., 2012. GIS-based assessment of landslide susceptibility on the base of the Weights-of-Evidence model. Landslides 9, 511–528. https://doi.org/10.1007/s10346-011-0305-5

Nguyen, Q.H., Ly, H.-B., Ho, L.S., Al-Ansari, N., Le, H.V., Tran, V.Q., Prakash, I., Pham, B.T., 2021. Influence of Data Splitting on Performance of Machine Learning Models in Prediction of Shear Strength of Soil. Mathematical Problems in Engineering 2021, 1–15. https://doi.org/10.1155/2021/4832864

Nhu, V.-H., Shirzadi, A., Shahabi, H., Chen, W., Clague, J.J., Geertsema, M., Jaafari, A., Avand, M., Miraki, S., Talebpour Asl, D., Pham, B.T., Ahmad, B.B., Lee, S., 2020. Shallow Landslide Susceptibility Mapping by Random Forest Base Classifier and Its Ensembles in a Semi-Arid Region of Iran. Forests 11. https://doi.org/10.3390/f11040421

Nichols Jr., T.C., 1980. Rebound, its nature and effect on engineering works. Quarterly Journal of Engineering Geology & Hydrogeology 13, 133–152.

Nigrelli, G., Fratianni, S., Zampollo, A., Turconi, L., Chiarle, M., 2018. The altitudinal temperature lapse rates applied to high elevation rockfalls studies in the Western European Alps. Theoretical and Applied Climatology 131, 1479–1491. https://doi.org/10.1007/s00704-017-2066-0

Nishii, R., Matsuoka, N., Daimaru, H., Yasuda, M., 2013. Precursors and triggers of an alpine rockslide in Japan: the 2004 partial collapse during a snow-melting period. Landslides 10, 75–82. https://doi.org/10.1007/s10346-012-0353-5

Nohani, E., Moharrami, M., Sharafi, S., Khosravi, K., Pradhan, B., Pham, B.T., Lee, S., M. Melesse, A., 2019. Landslide Susceptibility Mapping Using Different GIS-Based Bivariate Models. Water 11. https://doi.org/10.3390/w11071402

Notti, D., Davalillo, J.C., Herrera, G., Mora, O., 2010. Assessment of the performance of X-band satellite radar data for landslide mapping and monitoring: Upper Tena Valley case study. Natural Hazards and Earth System Sciences 10, 1865–1875. https://doi.org/10.5194/nhess-10-1865-2010

Notti, D., Herrera, G., Bianchini, S., Meisina, C., García-Davalillo, J.C., Zucca, F., 2014. A methodology for improving landslide PSI data analysis. null 35, 2186–2214. https://doi.org/10.1080/01431161.2014.889864

Nourani, Y., Andresen, B., 1998. A comparison of simulated annealing cooling strategies. Journal of Physics A: Mathematical and General 31, 8373–8385. https://doi.org/10.1088/0305-4470/31/41/011

Nunes, L.M., Caeiro, S., Cunha, M.C., Ribeiro, L., 2006. Optimal estuarine sediment monitoring network design with simulated annealing. Journal of Environmental Management 78, 294–304. https://doi.org/10.1016/j.jenvman.2005.04.024

Nussbaum, C., Marquer, D., Biino, G.G., 1998. Two subduction events in a polycyclic basement: Alpine and pre-Alpine high-pressure metamorphism in the Suretta nappe, Swiss Eastern Alps. Journal of Metamorphic Geology 16, 591–605. https://doi.org/10.1111/j.1525-1314.1998.00154.x

Nychka D., Furrer R., Paige J., Sain S., 2017. “fields: Tools for spatial data.” doi: 10.5065/D6W957CT, R package version 11.5.

Oberkampf, W.L., Helton, J.C., Joslyn, C.A., Wojtkiewicz, S.F., Ferson, S., 2004. Challenge problems: uncertainty in system response given uncertain parameters. Reliability Engineering & System Safety 85, 11–19. https://doi.org/10.1016/j.ress.2004.03.002

Occhiena, C., Coviello, V., Arattano, M., Chiarle, M., Morra di Cella, U., Pirulli, M., Pogliotti, P., Scavia, C., 2012. Analysis of microseismic signals and temperature recordings for rock slope stability investigations in high mountain areas. Natural Hazards and Earth System Sciences 12, 2283–2298. https://doi.org/10.5194/nhess-12-2283-2012

Öge, İ.F., 2017. Assessing Rock Mass Permeability Using Discontinuity Properties. Procedia Engineering 191, 638–645. https://doi.org/10.1016/j.proeng.2017.05.373

Oliver, M., Webster, R., Gerrard, J., 1989. Geostatistics in Physical Geography. Part I: Theory. Transactions of the Institute of British Geographers 14, 259. https://doi.org/10.2307/622687

Owens, H.L., Campbell, L.P., Dornak, L.L., Saupe, E.E., Barve, N., Soberón, J., Ingenloff, K., Lira-Noriega, A., Hensz, C.M., Myers, C.E., Peterson, A.T., 2013. Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. Ecological Modelling 263, 10–18. https://doi.org/10.1016/j.ecolmodel.2013.04.011

Ozturk, U., Pittore, M., Behling, R., Roessner, S., Andreani, L., Korup, O., 2021. How robust are landslide susceptibility estimates? Landslides 18, 681–695. https://doi.org/10.1007/s10346-020-01485-5

Palmstrom A., 1982. The volumetric joint count: a useful and simple measure of the degree of rock mass jointing. Proceedings of the IAEG congress New Delhi Vol. 2, 221–28.

Pantelidis, L., 2009. Rock slope stability assessment through rock mass classification systems. International Journal of Rock Mechanics and Mining Sciences 46, 315–325. https://doi.org/10.1016/j.ijrmms.2008.06.003

Pappalardo, G., 2018. First results of infrared thermography applied to the evaluation of hydraulic conductivity in rock masses. Hydrogeology Journal 26, 417–428. https://doi.org/10.1007/s10040-017-1670-5

Paranunzio, R., Chiarle, M., Laio, F., Nigrelli, G., Turconi, L., Luino, F., 2019. New insights in the relation between climate and slope failures at high-elevation sites. Theoretical and Applied Climatology 137, 1765–1784. https://doi.org/10.1007/s00704-018-2673-4

Paranunzio, R., Laio, F., Chiarle, M., Nigrelli, G., Guzzetti, F., 2016. Climate anomalies associated with the occurrence of rockfalls at high-elevation in the Italian Alps. Natural Hazards and Earth System Sciences 16, 2085–2106. https://doi.org/10.5194/nhess-16-2085-2016

Paranunzio, R., Laio, F., Nigrelli, G., Chiarle, M., 2015. A method to reveal climatic variables triggering slope failures at high elevation. Natural Hazards 76, 1039–1061. https://doi.org/10.1007/s11069-014-1532-6

Park, N. -W., Chi, K. -H., 2008. Quantitative assessment of landslide susceptibility using high-resolution remote sensing data and a generalized additive model. null 29, 247–264. https://doi.org/10.1080/01431160701227661

Parker, R.N., Densmore, A.L., Rosser, N.J., de Michele, M., Li, Y., Huang, R., Whadcoat, S., Petley, D.N., 2011. Mass wasting triggered by the 2008 Wenchuan earthquake is greater than orogenic growth. Nature Geoscience 4, 449–452. https://doi.org/10.1038/ngeo1154

Paudel, U., Oguchi, T., 2014. Implementation of Random Forest in landslide susceptibility study, a case study of the Tokamachi area, Niigata, Japan. Japan Geoscience Union.

Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. Computers & Geosciences 30, 683–691. https://doi.org/10.1016/j.cageo.2004.03.012

Pepin, N., Bradley, R.S., Diaz, H.F., Baraer, M., Caceres, E.B., Forsythe, N., Fowler, H., Greenwood, G., Hashmi, M.Z., Liu, X.D., Miller, J.R., Ning, L., Ohmura, A., Palazzi, E., Rangwala, I., Schöner, W., Severskiy, I., Shahgedanova, M., Wang, M.B., Williamson, S.N., Yang, D.Q., Mountain Research Initiative EDW Working Group, 2015. Elevation-dependent warming in mountain regions of the world. Nature Climate Change 5, 424–430. https://doi.org/10.1038/nclimate2563

Pereira, S., Zêzere, J.L., Bateira, C., 2012. Technical Note: Assessing predictive capacity and conditional independence of landslide predisposing factors for shallow landslide susceptibility models. Natural Hazards and Earth System Sciences 12, 979–988. https://doi.org/10.5194/nhess-12-979-2012

Persichillo, P., Bordoni, M., Meisina, C., Bartelletti, C., Giannecchini, R., D'Amato Avanzi, G., Galanti, Y., Cevasco, A., Brandolini, P., Galve, J., Barsanti, M., 2016. Shallow landslide susceptibility analysis in relation to land use scenarios. pp. 1605–1612. https://doi.org/10.1201/b21520-199

Peruccacci, S., Brunetti, M.T., Gariano, S.L., Melillo, M., Rossi, M., Guzzetti, F., 2017. Rainfall thresholds for possible landslide occurrence in Italy. Geomorphology 290, 39–57. https://doi.org/10.1016/j.geomorph.2017.03.031

Petschko H, Bell R, Leopold P, et al., 2013. Landslide inventories for reliable susceptibility maps in Lower Austria. In: Margottini C, Canuti P, Sassa K (eds) Landslide Science and Practice, Springer, pp 281–286

Petschko, H., Bell, R., Brenning, A., Glade, T., 2012. Landslide susceptibility modeling with generalized additive models–facing the heterogeneity of large regions. Landslides and Engineered Slopes, Protecting Society through Improved Understanding 1, 769–777.

Petschko, H., Bell, R., Glade, T., 2016. Effectiveness of visually analyzing LiDAR DTM derivatives for earth and debris slide inventory mapping for statistical susceptibility modeling. Landslides 13, 857–872. https://doi.org/10.1007/s10346-015-0622-1

Petschko, H., Brenning, A., Bell, R., Goetz, J., Glade, T., 2014. Assessing the quality of landslide susceptibility maps – case study Lower Austria. Natural Hazards and Earth System Sciences 14, 95–118. https://doi.org/10.5194/nhess-14-95-2014

Pham, B.T., Jaafari, A., Nguyen-Thoi, T., Van Phong, T., Nguyen, H.D., Satyam, N., Masroor, M., Rehman, S., Sajjad, H., Sahana, M., Van Le, H., Prakash, I., 2021. Ensemble machine learning models based on Reduced Error Pruning Tree for prediction of rainfall-induced landslides. null 14, 575–596. https://doi.org/10.1080/17538947.2020.1860145

Pham, B.T., Prakash, I., Dou, J., Singh, S.K., Trinh, P.T., Tran, H.T., Le, T.M., Van Phong, T., Khoi, D.K., Shirzadi, A., Bui, D.T., 2020. A novel hybrid approach of landslide susceptibility modelling using rotation forest ensemble and different base classifiers. null 35, 1267–1292. https://doi.org/10.1080/10106049.2018.1559885

Pham, B.T., Prakash, I., Singh, S.K., Shirzadi, A., Shahabi, H., Tran, T.-T.-T., Bui, D.T., 2019. Landslide susceptibility modeling using Reduced Error Pruning Trees and different ensemble techniques: Hybrid machine learning approaches. CATENA 175, 203–218. https://doi.org/10.1016/j.catena.2018.12.018

Pignone, F. and Rebora, N., 2014: GRISO: Rainfall Generator of Spatial Interpolation from Observation, Geophys.Res. Abstr. EGU Gen. Assem. 2014, 16, EGU2014-13946.

Pinheiro, M., Emery, X., Miranda, T., Lamas, L., Espada, M., 2018. Modelling Geotechnical Heterogeneities Using Geostatistical Simulation and Finite Differences Analysis. Minerals 8. https://doi.org/10.3390/min8020052

Pinheiro, M., Vallejos, J., Miranda, T., Emery, X., 2016. Geostatistical simulation to map the spatial heterogeneity of geomechanical parameters: A case study with rock mass rating. Engineering Geology 205, 93–103. https://doi.org/10.1016/j.enggeo.2016.03.003

Pistocchi, A., 2016. Simple estimation of snow density in an Alpine region. Journal of Hydrology: Regional Studies 6, 82–89. https://doi.org/10.1016/j.ejrh.2016.03.004

Ponziani, M., Pogliotti, P., Stevenin, H., Ratto, S.M., 2020. Debris-flow Indicator for an early warning system in the Aosta valley region. Natural Hazards 104, 1819–1839. https://doi.org/10.1007/s11069-020-04249-5

Pourghasemi, H.R., Rahmati, O., 2018. Prediction of the landslide susceptibility: Which algorithm, which precision? CATENA 162, 177–192. https://doi.org/10.1016/j.catena.2017.11.022

Price, J., 2016. Implications of groundwater behaviour on the geomechanics of rock slope stability, in: Proceedings of the First Asia Pacific Slope Stability in Mining Conference. Presented at the First Asia Pacific Slope Stability in Mining Conference, Australian Centre for Geomechanics, Perth, pp. 25–48. https://doi.org/10.36487/ACG_rep/1604_0.3_Price

Probst, P., Boulesteix, A.-L., 2017. To tune or not to tune the number of trees in random forest? Journal of Machine Learning Research 18.

Probst, P., Wright, M.N., Boulesteix, A., 2019. Hyperparameters and tuning strategies for random forest. WIREs Data Mining Knowl Discov 9. https://doi.org/10.1002/widm.1301

Pya, N., Wood, S.N., 2016. A note on basis dimension selection in generalized additive modelling. arXiv:1602.06696.

Qin, C.-Z., Bao, L.-L., Zhu, A.-X., Wang, R.-X., Hu, X.-M., 2013. Uncertainty due to DEM error in landslide susceptibility mapping. International Journal of Geographical Information Science 27, 1364–1380. https://doi.org/10.1080/13658816.2013.770515

Rabby, Y.W., Ishtiaque, A., Rahman, Md.S., 2020a. Evaluating the Effects of Digital Elevation Models in Landslide Susceptibility Mapping in Rangamati District, Bangladesh. Remote Sensing 12. https://doi.org/10.3390/rs12172718

Rabby, Y.W., Li, Y., 2020b. Landslide Susceptibility Mapping Using Integrated Methods: A Case Study in the Chittagong Hilly Areas, Bangladesh. Geosciences 10. https://doi.org/10.3390/geosciences10120483

Raffa, M., 2020. Il ruolo della neve nell'innesco di frane superficiali valutato con random forest: il caso del Mont Cervin e del Mont Emilius in Valle D'Aosta (Unpublished Master Thesis). Università degli Studi di Milano.

Rajczak, J., Pall, P., Schär, C., 2013. Projections of extreme precipitation events in regional climate simulations for Europe and the Alpine Region. Journal of Geophysical Research: Atmospheres 118, 3610–3626. https://doi.org/10.1002/jgrd.50297

Raspini, F., Bardi, F., Bianchini, S., Ciampalini, A., Del Ventisette, C., Farina, P., Ferrigno, F., Solari, L., Casagli, N., 2017. The contribution of satellite SAR-derived displacement measurements in landslide risk management practices. Natural Hazards 86, 327–351. https://doi.org/10.1007/s11069-016-2691-4

Ratto, S., Bonetto, F., Comoglio, C., 2003. The October 2000 flooding in Valle d'Aosta (Italy): Event description and land planning measures for the risk mitigation. International Journal of River Basin Management 1, 105–116. https://doi.org/10.1080/15715124.2003.9635197

Ravanel, L., Deline, P., 2010. Climate influence on rockfalls in high-Alpine steep rockwalls: The north side of the Aiguilles de Chamonix (Mont Blanc massif) since the end of the 'Little Ice Age.' The Holocene 21, 357–365. https://doi.org/10.1177/0959683610374887

Ravanel, L., Deline, P., Lambiel, C., Vincent, C., 2013. Instability of a high alpine rock ridge: the Lower Arête Des Cosmiques, Mont Blanc Massif, France. Geografiska Annaler. Series A, Physical Geography 95, 51–66.

Reichenbach, P., Busca, C., Mondini, A.C., Rossi, M., 2014. The Influence of Land Use Change on Landslide Susceptibility Zonation: The Briga Catchment Test Site (Messina, Italy). Environmental Management 54, 1372–1384. https://doi.org/10.1007/s00267-014-0357-0

Reichenbach, P., Rossi, M., Malamud, B.D., Mihir, M., Guzzetti, F., 2018. A review of statistically-based landslide susceptibility models. Earth-Science Reviews 180, 60–91. https://doi.org/10.1016/j.earscirev.2018.03.001

Remondo, J., González, A., De Terán, J.R.D., Cendrero, A., Fabbri, A., Chung, C.-J.F., 2003. Validation of Landslide Susceptibility Maps; Examples and Applications from a Case Study in Northern Spain. Natural Hazards 30, 437–449. https://doi.org/10.1023/B:NHAZ.0000007201.80743.fc

Reygondeau, G., Guieu, C., Benedetti, F., Irisson, J.-O., Ayata, S.-D., Gasparini, S., Koubbi, P., 2017. Biogeochemical regions of the Mediterranean Sea: An objective multidimensional and multivariate environmental approach. Progress in Oceanography 151, 138–148. https://doi.org/10.1016/j.pocean.2016.11.001

Riva, F., Agliardi, F., Amitrano, D., Crosta, G.B., 2018. Damage-Based Time-Dependent Modeling of Paraglacial to Postglacial Progressive Failure of Large Rock Slopes. Journal of Geophysical Research: Earth Surface 123, 124–141. https://doi.org/10.1002/2017JF004423

Roger Bivand and Danlin Yu (2020). spgwr: Geographically Weighted Regression. R package version 0.6-34

Rolland, C., 2003. Spatial and Seasonal Variations of Air Temperature Lapse Rates in Alpine Regions. Journal of Climate 16, 1032–1046. https://doi.org/10.1175/1520-0442(2003)016<1032:SASVOA>2.0.CO;2

Rosi, A., Tofani, V., Tanteri, L., Tacconi Stefanelli, C., Agostini, A., Catani, F., Casagli, N., 2018. The new landslide inventory of Tuscany (Italy) updated with PS-InSAR: geomorphological features and landslide distribution. Landslides 15, 5–19. https://doi.org/10.1007/s10346-017-0861-4

Rosin P.L., Hervás, J., 2005. Remote sensing image thresholding methods for determining landslide activity. International Journal of Remote Sensing 26, 1075–1092. https://doi.org/10.1080/01431160512331330481

Rossi, M., Guzzetti, F., Reichenbach, P., Mondini, A.C., Peruccacci, S., 2010. Optimal landslide susceptibility zonation based on multiple forecasts. Geomorphology 114, 129–142. https://doi.org/10.1016/j.geomorph.2009.06.020

Rossi, M., Reichenbach, P., 2016. LAND-SE: a software for statistically based landslide susceptibilityzonation, version 1.0. Geosci. Model Dev. 9, 3533–3543. https://doi.org/10.5194/gmd-9-3533-2016

Rossi, M., Sarro, R., Reichenbach, P., Mateos, R.M., 2021. Probabilistic identification of rockfall source areas at regional scale in El Hierro (Canary Islands, Spain). Geomorphology 381, 107661. https://doi.org/10.1016/j.geomorph.2021.107661

Rotigliano, E., Agnesi, V., Cappadonia, C., Conoscenti, C., 2011. The role of the diagnostic areas in the assessment of landslide susceptibility models: a test in the sicilian chain. Nat Hazards 58, 981–999. https://doi.org/10.1007/s11069-010-9708-1

Rotigliano, E., Cappadonia, C., Conoscenti, C., Costanzo, D., Agnesi, V., 2012. Slope units-based flow susceptibility model: using validation tests to select controlling factors. Natural Hazards 61, 143–153. https://doi.org/10.1007/s11069-011-9846-0

Roy, C.J., Oberkampf, W.L., 2011. A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing. Computer Methods in Applied Mechanics and Engineering 200, 2131–2144.

Roy, J., Saha, S., Arabameri, A., Blaschke, T., Bui, D.T., 2019. A Novel Ensemble Approach for Landslide Susceptibility Mapping (LSM) in Darjeeling and Kalimpong Districts, West Bengal, India. Remote Sensing 11. https://doi.org/10.3390/rs11232866

Rubel F, Brugger K, Haslinger K, Auer I., 2017. The climate of the European Alps: Shift of very high resolution Köppen-Geiger climate zones 1800–2100. Meteorologische Zeitschrift 26(2):115–125

Ruff, M., Czurda, K., 2008. Landslide susceptibility analysis with a heuristic approach in the Eastern Alps (Vorarlberg, Austria). Geomorphology 94, 314–324. https://doi.org/10.1016/j.geomorph.2006.10.032

Ruß, G., Brenning, A., 2010. Data Mining in Precision Agriculture: Management of Spatial Information, in: Hüllermeier, E., Kruse, R., Hoffmann, F. (Eds.), Computational Intelligence for Knowledge-Based Systems Design. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 350–359.

Sameen, M.I., Pradhan, B., Bui, D.T., Alamri, A.M., 2020. Systematic sample subdividing strategy for training landslide susceptibility models. CATENA 187, 104358. https://doi.org/10.1016/j.catena.2019.104358

Samia, J., Temme, A., Bregt, A., Wallinga, J., Guzzetti, F., Ardizzone, F., Rossi, M., 2017a. Do landslides follow landslides? Insights in path dependency from a multi-temporal landslide inventory. Landslides 14, 547–558.

Samia, J., Temme, A., Bregt, A., Wallinga, J., Guzzetti, F., Ardizzone, F., Rossi, M., 2017b. Characterization and quantification of path dependency in landslide susceptibility. Geomorphology 292, 16–24. https://doi.org/10.1016/j.geomorph.2017.04.039

Samia, J., Temme, A., Bregt, A., Wallinga, J., Guzzetti, F., Ardizzone, F., 2020. Dynamic path-dependent landslide susceptibility modelling. Nat. Hazards Earth Syst. Sci. 20, 271–285. https://doi.org/10.5194/nhess-20-271-2020

Samia, J., Temme, A., Bregt, A.K., Wallinga, J., Stuiver, J., Guzzetti, F., Ardizzone, F., Rossi, M., 2018. Implementing landslide path dependency in landslide susceptibility modelling. Landslides 15, 2129–2144. https://doi.org/10.1007/s10346-018-1024-y

Santangelo, M., Cardinali, M., Rossi, M., Mondini, A.C., Guzzetti, F., 2010. Remote landslide mapping using a laser rangefinder binocular and GPS. Natural Hazards and Earth System Sciences 10, 2539–2546. https://doi.org/10.5194/nhess-10-2539-2010

Santangelo, M., Marchesini, I., Cardinali, M., Fiorucci, F., Rossi, M., Bucci, F., Guzzetti, F., 2015. A method for the assessment of the influence of bedding on landslide abundance and types. Landslides 12, 295–309. https://doi.org/10.1007/s10346-014-0485-x

Sarout, J., Cazes, E., Delle Piane, C., Arena, A., Esteban, L., 2017. Stress-dependent permeability and wave dispersion in tight cracked rocks: Experimental validation of simple effective medium models. Journal of Geophysical Research: Solid Earth 122, 6180–6201. https://doi.org/10.1002/2017JB014147

Scaioni, M., Longoni, L., Melillo, V., Papini, M., 2014. Remote Sensing for Landslide Investigations: An Overview of Recent Achievements and Perspectives. Remote Sensing 6, 9600–9652. https://doi.org/10.3390/rs6109600

Scavia, C., Barbero, M., Castelli, M., Marchelli, M., Peila, D., Torsello, G., Vallero, G., 2020. Evaluating Rockfall Risk: Some Critical Aspects. Geosciences 10. https://doi.org/10.3390/geosciences10030098

Scheiber, T., Pfiffner, O.A., Schreurs, G., 2012. Strain accumulation during basal accretion in continental collision — A case study from the Suretta nappe (eastern Swiss Alps). Tectonophysics 579, 56–73. https://doi.org/10.1016/j.tecto.2012.03.009

Schilirò, L., Cepeda, J., Devoli, G., Piciullo, L., 2021. Regional Analyses of Rainfall-Induced Landslide Initiation in Upper Gudbrandsdalen (South-Eastern Norway) Using TRIGRS Model. Geosciences 11, 35. https://doi.org/10.3390/geosciences11010035

Schuster, R.L., Highland, L.M., 2007. The Third Hans Cloos Lecture. Urban landslides: socioeconomic impacts and overview of mitigative strategies. Bulletin of Engineering Geology and the Environment 66, 1–27. https://doi.org/10.1007/s10064-006-0080-z

Scott, D.N., Wohl, E.E., 2019. Bedrock fracture influences on geomorphic process and form across process domains and scales: Bedrock Fracture Influences on Geomorphology. Earth Surf. Process. Landforms 44, 27–45. https://doi.org/10.1002/esp.4473

Scudiero, E., Lesch, S.M., Corwin, D.L., 2016. Validation of Sensor-Directed Spatial Simulated Annealing Soil Sampling Strategy. J. Environ. Qual. 45, 1226–1233. https://doi.org/10.2134/jeq2015.09.0458

Segal, M. R., 2004. Machine Learning Benchmarks and Random Forest Regression. UCSF: Center for Bioinformatics and Molecular Biostatistics. Retrieved from https://escholarship.org/uc/item/35x3v9t4

Segoni, S., Lagomarsino, D., Fanti, R., Moretti, S., Casagli, N., 2015. Integration of rainfall thresholds and susceptibility maps in the Emilia Romagna (Italy) regional-scale landslide warning system. Landslides 12, 773–785. https://doi.org/10.1007/s10346-014-0502-0

Segoni, S., Pappafico, G., Luti, T., Catani, F., 2020. Landslide susceptibility assessment in complex geological settings: sensitivity to geological information and insights on its parameterization. Landslides 17, 2443–2453. https://doi.org/10.1007/s10346-019-01340-2

Segoni, S., Tofani, V., Rosi, A., Catani, F., Casagli, N., 2018. Combination of Rainfall Thresholds and Susceptibility Maps for Dynamic Landslide Hazard Assessment at Regional Scale. Frontiers in Earth Science 6, 85. https://doi.org/10.3389/feart.2018.00085

Seibert, J., Jenicek, M., Huss, M., Ewen, T., 2015. Chapter 4 - Snow and Ice in the Hydrosphere, in: Shroder, J.F., Haeberli, W., Whiteman, C. (Eds.), Snow and Ice-Related Hazards, Risks and Disasters. Academic Press, Boston, pp. 99–137. https://doi.org/10.1016/B978-0-12-394849-6.00004-4

Seibold, H., Bernau, C., Boulesteix, A.-L., De Bin, R., 2018. On the choice and influence of the number of boosting steps for high-dimensional linear Cox-models. Computational Statistics 33, 1195–1215. https://doi.org/10.1007/s00180-017-0773-8

Selby MJ. 1982. Controls on the stability and inclinations of hillslopes formed on hard rock. Earth Surface Processes and Landforms 7:449–467.

Selby, M.J. 1982. Hillslope materials and processes. New York: Oxford University Press

Selvadurai, A.P.S., Boulon, M.J., Nguyen, T.S., 2005. The Permeability of an Intact Granite. pure and applied geophysics 162, 373–407. https://doi.org/10.1007/s00024-004-2606-2

Şen, Z., 2015. Chapter 2 - Basic Porous Medium Concepts, in: Şen, Z. (Ed.), Practical and Applied Hydrogeology. Elsevier, Oxford, pp. 43–97. https://doi.org/10.1016/B978-0-12-800075-5.00002-9

Sfondrini, G., Pasquarè, G., 2011 La Valchiavenna: un bacino pilota per il controllo dell'ambiente alpino";volume a cura di: Barbara Aldighieri, Guido Mazzoleni, Quaderni di Geodinamica Alpina - vol. N°10 pp. 243-257 , ISBN978-88-86596-12-1.

Shapley, L. S., 1953. A value for n-person games. In A. E. Roth (Ed.), The Shapley Value. Essays in Honor of Lloyd S. Shapley. (pp. 31–40). https://doi.org/10.2307/2554979

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

Shmueli, G., 2010. To explain or to predict? Statistical Science, 25, 289–310.

Shou, K.-J., Yang, C.-M., 2015. Predictive analysis of landslide susceptibility under climate change conditions — A study on the Chingshui River Watershed of Taiwan. Engineering Geology 192, 46–62. https://doi.org/10.1016/j.enggeo.2015.03.012

Simpson, G.L., 2018. Modelling Palaeoecological Time Series Using Generalised Additive Models. Front. Ecol. Evol. 6, 149. https://doi.org/10.3389/fevo.2018.00149

Snow, D.T., 1968. Rock Fracture Spacings, Openings, and Porosities. Journal of the Soil Mechanics and Foundations Division 94, 73–91. https://doi.org/10.1061/JSFEAQ.0001097

Sperl, J., Trckova, J., 2008. Permeability and porosity of rocks and their relationship based on laboratory testing. Acta Geodynamica et Geomaterialia 5, 41–47.

Stahl, T., Tye, A., 2020. Schmidt hammer and terrestrial laser scanning (TLS) used to detect single-event displacements on the Pleasant Valley fault (Nevada, USA). Earth Surface Processes and Landforms 45, 473–483.

Stahl, T., Winkler, S., Quigley, M., Bebbington, M., Duffy, B., Duke, D., 2013. Schmidt hammer exposure-age dating (SHD) of late Quaternary fluvial terraces in New Zealand. Earth Surf. Process. Landforms 38, 1838–1850. https://doi.org/10.1002/esp.3427

Steger, S., Brenning, A., Bell, R., Glade, T., 2016a. The propagation of inventory-based positional errors into statistical landslide susceptibility models. Natural Hazards and Earth System Sciences 16, 2729–2745. https://doi.org/10.5194/nhess-16-2729-2016

Steger, S., Brenning, A., Bell, R., Glade, T., 2017. The influence of systematically incomplete shallow landslide inventories on statistical susceptibility models and suggestions for improvements. Landslides 14, 1767–1781. https://doi.org/10.1007/s10346-017-0820-0

Steger, S., Brenning, A., Bell, R., Petschko, H., Glade, T., 2016b. Exploring discrepancies between quantitative validation results and the geomorphic plausibility of statistical landslide susceptibility maps. Geomorphology 262, 8–23. https://doi.org/10.1016/j.geomorph.2016.03.015

Steger, S., Mair, V., Kofler, C., Pittore, M., Zebisch, M., Schneiderbauer, S., 2021. Correlation does not imply geomorphic causation in data-driven landslide susceptibility modelling – Benefits of exploring landslide data collection effects. Science of The Total Environment 776, 145935. https://doi.org/10.1016/j.scitotenv.2021.145935

Sterlacchini, S., Ballabio, C., Blahut, J., Masetti, M., Sorichetta, A., 2011. Spatial agreement of predicted patterns in landslide susceptibility maps. Geomorphology 125, 51–61. https://doi.org/10.1016/j.geomorph.2010.09.004

Stoffel, M., Ballesteros Cánovas, J.A., Luckman, B.H., Casteller, A., Villalba, R., 2019. Tree-ring correlations suggest links between moderate earthquakes and distant rockfalls in the Patagonian Cordillera. Scientific Reports 9, 12112. https://doi.org/10.1038/s41598-019-48530-5

Stoffel, M., Tiranti, D., Huggel, C., 2014. Climate change impacts on mass movements — Case studies from the European Alps. Science of The Total Environment 493, 1255–1266. https://doi.org/10.1016/j.scitotenv.2014.02.102

Strobl, C., Boulesteix, A.-L., Zeileis, A., Hothorn, T., 2007. Bias in random forest variable importance measures: Illustrations, sources and a solution. BMC Bioinformatics 8, 25. https://doi.org/10.1186/1471-2105-8-25

Strozzi, T., Wegmuller, U., Keusen, H.R., Graf, K., Wiesmann, A., 2006. Analysis of the terrain displacement along a funicular by SAR interferometry. IEEE Geoscience and Remote Sensing Letters 3, 15–18. https://doi.org/10.1109/LGRS.2005.855072

Strunden, J., Ehlers, T.A., Brehm, D., Nettesheim, M., 2015. Spatial and temporal variations in rockfall determined from TLS measurements in a deglaciated valley, Switzerland. Journal of Geophysical Research: Earth Surface 120, 1251–1273. https://doi.org/10.1002/2014JF003274

Stumpf, A., Kerle, N., 2011. Object-oriented mapping of landslides using Random Forests. Remote Sensing of Environment 115, 2564–2577. https://doi.org/10.1016/j.rse.2011.05.013

Stumvoll, M.J., Canli, E., Engels, A., Thiebes, B., Groiss, B., Glade, T., Schweigl, J., Bertagnoli, M., 2020. The "Salcher" landslide observatory—experimental long-term monitoring in the Flysch Zone of Lower Austria. Bulletin of Engineering Geology and the Environment 79, 1831–1848. https://doi.org/10.1007/s10064-019-01632-w

Subramanian, S., Fan, X., Yunus, A.P., van Asch, T., Scaringi, G., Xu, Q., Dai, L., Ishikawa, T., Huang, R., 2020. A Sequentially Coupled Catchment-Scale Numerical Model for Snowmelt-Induced Soil Slope Instabilities. Journal of Geophysical Research: Earth Surface 125, e2019JF005468. https://doi.org/10.1029/2019JF005468

Tantardini, D., 2016. Geologia del quaternario e geomorfologia della bassa Valchiavenna (SO). PhD Thesis. Università degli Studi di Milano.

Tantardini, D., Riganti, N., Taglieri, P., De Finis, E., Bini, A., 2013. Glacier dynamics in San Giacomo Valley (Central Alps, Sondrio, Italy).

Tanyas, H., Hill, K., Mahoney, L., Fadel, I., Lombardo, L., 2021. The world's second-largest, recorded landslide event: lessons learnt from the landslides triggered during and after the 2018 Mw 7.5 Papua New Guinea earthquake. EarthArXiv, 1-46. https://doi.org/10.31223/X5KS5H

Tarolli, P., 2014. High-resolution topography for understanding Earth surface processes: Opportunities and challenges. Geomorphology 216, 295–312.

Tartarotti, P., Guerini, S., Rotondo, F., Festa, A., Balestro, G., Bebout, G.E., Cannaò, E., Epstein, G.S., Scambelluri, M., 2019. Superposed Sedimentary and Tectonic Block-In-Matrix Fabrics in a Subducted Serpentinite Mélange (High-Pressure Zermatt Saas Ophiolite, Western Alps). Geosciences 9. https://doi.org/10.3390/geosciences9080358

Temme, A., Guzzetti, F., Samia, J., Mirus, B.B., 2020. The future of landslides' past—a framework for assessing consecutive landsliding systems. Landslides 17, 1519–1528. https://doi.org/10.1007/s10346-020-01405-7

Thiery, Y., Malet, J.-P., Sterlacchini, S., Puissant, A., Maquaire, O., 2007. Landslide susceptibility assessment by bivariate methods at large scales: Application to a complex mountainous environment. Geomorphology 92, 38–59. https://doi.org/10.1016/j.geomorph.2007.02.020

Tibaldi, A., Pasquarè, F.A., 2008. Quaternary deformations along the 'Engadine–Gruf tectonic system', Swiss–Italian Alps. Journal of Quaternary Science 23, 475–487. https://doi.org/10.1002/jqs.1150

Torres-Suarez, M.C., Alarcon-Guzman, A., Moya, R., Berdugo-D., 2014. Effects of loading–unloading and wetting–drying cycles on geomechanical behaviors of mudrocks in the Colombian Andes. Journal of Rock Mechanics and Geotechnical Engineering 6, 257–268. https://doi.org/10.1016/j.jrmge.2014.04.004

Touw, W.G., Bayjanov, J.R., Overmars, L., Backus, L., Boekhorst, J., Wels, M., van Hijum, S.A.F.T., 2013. Data mining in the Life Sciences with Random Forest: a walk in the park or lost in the jungle? Briefings in Bioinformatics 14, 315–326. https://doi.org/10.1093/bib/bbs034

Triglia, A., Frattini, P., Casagli, N., Catani, F., Crosta, G., Esposito, C., Iadanza, C., Lagomarsino, D., Mugnozza, G.S., Segoni, S., 2013. Landslide susceptibility mapping at national scale: the Italian case study, in: Landslide Science and Practice. Springer, pp. 287–295.

Triki, E., Collette, Y., Siarry, P., 2005. A theoretical study on the behavior of simulated annealing leading to a new cooling schedule. European Journal of Operational Research 166, 77–92. https://doi.org/10.1016/j.ejor.2004.03.035

Umar, Z., Pradhan, B., Ahmad, A., Jebur, M.N., Tehrany, M.S., 2014. Earthquake induced landslide susceptibility mapping using an integrated ensemble frequency ratio and logistic regression models in West Sumatera Province, Indonesia. CATENA 118, 124–135. https://doi.org/10.1016/j.catena.2014.02.005

USDA-NEH, 2004. National Engineering Handbook Chapter 11 Part 630 United States Department of Agriculture Natural Resources Conservation Service

Vakhshoori, V., Zare, M., 2016. Landslide susceptibility mapping by comparing weight of evidence, fuzzy logic, and frequency ratio methods. Geomatics, Natural Hazards and Risk 7, 1731–1752. https://doi.org/10.1080/19475705.2016.1144655

Valagussa, A., Frattini, P., Crosta, G.B., 2014. Earthquake-induced rockfall hazard zoning. Engineering Geology 182, 213–225. https://doi.org/10.1016/j.enggeo.2014.07.009

Van Den Eeckhaut, M., Kerle, N., Poesen, J., Hervás, J., 2012. Object-oriented identification of forested landslides with derivatives of single pulse LiDAR data. Geomorphology 173–174, 30–42. https://doi.org/10.1016/j.geomorph.2012.05.024

Van Den Eeckhaut, M., Reichenbach, P., Guzzetti, F., Rossi, M., Poesen, J., 2009. Combined landslide inventory and susceptibility assessment based on different mapping units: an example from the Flemish

Van der Hoven, S.J., Solomon, D.K., Moline, G.R., 2003. Modeling unsaturated flow and transport in the saprolite of fractured sedimentary rocks: Effects of periodic wetting and drying. Water Resources Research 39. https://doi.org/10.1029/2002WR001926

Van Groenigen, J.W., Pieters, G., Stein, A., 2000. Optimizing spatial sampling for multivariate contamination in urban areas. Environmetrics 11, 227–244.

Van Groenigen, J.W., Siderius, W., Stein, A., 1999. Constrained optimisation of soil sampling for minimisation of the kriging variance. Geoderma 87, 239–259. https://doi.org/10.1016/S0016-7061(98)00056-1

Van Groenigen, J.W., Stein, A., 1998. Constrained optimization of spatial sampling using continuous simulated annealing. J. Environ. Qual. 27, 1078–1086.

Van Westen, C.J., Castellanos, E., Kuriakose, S.L., 2008. Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. Engineering Geology 102, 112–131. https://doi.org/10.1016/j.enggeo.2008.03.010

Vargas, E. A., Velloso, R.Q., Chávez, L.E., Gusmão, L., do Amaral, C.P., 2013. On the Effect of Thermally Induced Stresses in Failures of Some Rock Slopes in Rio de Janeiro, Brazil. Rock Mechanics and Rock Engineering 46, 123–134. https://doi.org/10.1007/s00603-012-0247-9

Varnes D.J., 1978. Slope movement types and processes. In: Schuster RL, Krizek RJ (eds) Landslide analysis and control. Transportation Research Board, Special report 176. National Academy Sciences, Washington, DC, pp 11–33

Veronesi, F., Schillaci, C., 2019. Comparison between geostatistical and machine learning models as predictors of topsoil organic carbon with a focus on local uncertainty estimation. Ecological Indicators 101, 1032–1044. https://doi.org/10.1016/j.ecolind.2019.02.026

Viles, H.A., 2013. Linking weathering and rock slope instability: non-linear perspectives. Earth Surface Processes and Landforms 38, 62–70. https://doi.org/10.1002/esp.3294

Volkwein, A., Schellenberg, K., Labiouse, V., Agliardi, F., Berger, F., Bourrier, F., Dorren, L.K.A., Gerber, W., Jaboyedoff, M., 2011. Rockfall characterisation and structural protection – a review. Nat. Hazards Earth Syst. Sci. 11, 2617–2651. https://doi.org/10.5194/nhess-11-2617-2011

Voltolina D., 2021. Wildland surface fire behaviour: a spatial simulation model for operational emergency management. PhD Thesis, Università degli Studi di Milano. http://hdl.handle.net/2434/848088

Vorkauf, M., Marty, C., Kahmen, A., Hiltbrunner, E., 2021. Past and future snowmelt trends in the Swiss Alps: the role of temperature and snowpack. Climatic Change 165, 44. https://doi.org/10.1007/s10584-021-03027-x

Wahba, G., Wendelberger, J., 1980. Some New Mathematical Methods for Variational Objective Analysis Using Splines and Cross Validation. Mon. Wea. Rev. 108, 1122–1143. https://doi.org/10.1175/1520-0493(1980)108<1122:SNMMFV>2.0.CO;2

Walter, M., Schwaderer, U., Joswig, M., 2012. Seismic monitoring of precursory fracture signals from a destructive rockfall in the Vorarlberg Alps, Austria. Nat. Hazards Earth Syst. Sci. 12, 3545–3555. https://doi.org/10.5194/nhess-12-3545-2012

Wang, J., Wang, X., Zhou, S., Wu, S., Zhu, Y., Lu, C., 2016. Optimization of Sample Points for Monitoring Arable Land Quality by Simulated Annealing while Considering Spatial Variations. IJERPH 13, 980. https://doi.org/10.3390/ijerph13100980

Wang, J.-F., Jiang, C.-S., Hu, M.-G., Cao, Z.-D., Guo, Y.-S., Li, L.-F., Liu, T.-J., Meng, B., 2013. Design-based spatial sampling: Theory and implementation. Environmental Modelling & Software 40, 280–288. https://doi.org/10.1016/j.envsoft.2012.09.015

Wang, J.-F., Stein, A., Gao, B.-B., Ge, Y., 2012. A review of spatial sampling. Spatial Statistics 2, 1–14. https://doi.org/10.1016/j.spasta.2012.08.001

Wang, S., Ni, P., 2014. Application of block theory modeling on spatial block topological identification to rock slope stability analysis. International Journal of Computational Methods 11, 1350044. https://doi.org/10.1142/S0219876213500448

Wang, X., Crosta, G.B., Clague, J.J., Stead, D., Sun, J., Qi, S., Liu, H., 2021. Fault controls on spatial variation of fracture density and rock mass strength within the Yarlung Tsangpo Fault damage zone (southeastern Tibet). Engineering Geology 291, 106238. https://doi.org/10.1016/j.enggeo.2021.106238

Wang, Y., Fang, Z., Hong, H., 2019. Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China. Science of The Total Environment 666, 975–993. https://doi.org/10.1016/j.scitotenv.2019.02.263

Wasowski, J., Keefer, D.K., Lee, C.-T., 2011. Toward the next generation of research on earthquake-induced landslides: Current issues and future challenges. Engineering Geology 122, 1–8. https://doi.org/10.1016/j.enggeo.2011.06.001

Webster, R. and Oliver, M., 2001. Geostatistics for Environmental Scientists. John Wiley & Sons, Ltd, Chichester, 271 pp.

Webster, R., Oliver, M.A., 2007. Geostatistics for environmental scientists. John Wiley & Sons.

Wei, L.-W., Chen, H., Lee, C.-F., Huang, W.-K., Lin, M.-L., Chi, C.-C., Lin, H.-H., 2014. The mechanism of rockfall disaster: A case study from Badouzih, Keelung, in northern Taiwan. Engineering Geology 183, 116–126. https://doi.org/10.1016/j.enggeo.2014.10.008

Wenk HR., 1984. Brittle-ductile transition in the Northern Bergell Alps. Geologische Rundschau 73, 419–431.

Wiederkehr M., Bousquet R., Schmid S. & Berger A., 2008. From subduction to collision: Thermal overprint of HP/LT meta-sediments in the north-eastern Lepontine Dome (Swiss Alps) and consequences regarding the tectono-metamorphic evolution of the Alpine orogenic wedge. Swiss Journal of Geosciences, Vol.101, 127-155

Wilson RC, Jayko AS, 1997. Preliminary maps showing rainfall thresholds for debris-flow activity, San Francisco Bay Region, California. US Geological Survey Open-File Report 97-745 F

Wood, S.N., 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models: Estimation of Semiparametric Generalized Linear Models. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 73, 3–36. https://doi.org/10.1111/j.1467-9868.2010.00749.x

Wood, S.N., 2017. Generalized Additive Models: An Introduction with R, 2nd ed.; Chapman and Hall/CRC: Boca Raton, FL, USA.

Wood, S.N., Pya, N., Säfken, B., 2016. Smoothing Parameter and Model Selection for General Smooth Models. Journal of the American Statistical Association 111, 1548–1563. https://doi.org/10.1080/01621459.2016.1180986

World Meteorological Organization, 1989: Calculation of Monthly and Annual 30-Year Standard Normals. WCDP-No. 10, WMO-TD/No. 341, World Meteorological Organization.

World Meteorological Organization, 2007: The Role of Climatological Normals in a Changing Climate. WCDMP-No. 61, WMO-TD/No. 1377, World Meteorological Organization.

Wright, M.N., Ziegler, A., 2017. ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. Journal of Statistical Software 77, 1–17. https://doi.org/10.18637/jss.v077.i01

Yang, X., Chen, L., 2010. Using multi-temporal remote sensor imagery to detect earthquake-triggered landslides. International Journal of Applied Earth Observation and Geoinformation 12, 487–495. https://doi.org/10.1016/j.jag.2010.05.006

Yang, X., Wang, J., Hou, D., Zhu, C., He, M., 2018. Effect of Dry-Wet Cycling on the Mechanical Properties of Rocks: A Laboratory-Scale Experimental Study. Processes 6. https://doi.org/10.3390/pr6100199

Yang, X., Wang, J., Zhu, C., He, M., Gao, Y., 2019. Effect of wetting and drying cycles on microstructure of rock based on SEM. Environmental Earth Sciences 78, 183. https://doi.org/10.1007/s12665-019-8191-6

Yi, Y., Zhang, Z., Zhang, W., Jia, H., Zhang, J., 2020. Landslide susceptibility mapping using multiscale sampling strategy and convolutional neural network: A case study in Jiuzhaigou region. CATENA 195, 104851. https://doi.org/10.1016/j.catena.2020.104851

Youssef, A.M., Pourghasemi, H.R., Pourtaghi, Z.S., Al-Katheeri, M.M., 2016. Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia. Landslides 13, 839–856. https://doi.org/10.1007/s10346-015-0614-1

Youssef, A.M., Pradhan, B., Jebur, M.N., El-Harbi, H.M., 2015. Landslide susceptibility mapping using ensemble bivariate and multivariate statistical models in Fayfa area, Saudi Arabia. Environ Earth Sci 73, 3745–3761. https://doi.org/10.1007/s12665-014-3661-3

Zaniewski, A.E., Lehmann, A., Overton, J.M., 2002. Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. Ecological Modelling 157, 261–280. https://doi.org/10.1016/S0304-3800(02)00199-0

Zêzere, J., Neves, M., Henriques, C.S., Garcia, R., Oliveira, S., Piedade, A., 2009. Effects of landslide inventories uncertainty on landslide susceptibility modelling. Landslide processes: from geomorphologic mapping to dynamic modelling 81–86.

Zhang, L., 2013. Aspects of rock permeability. Frontiers of Structural and Civil Engineering 7, 102–116. https://doi.org/10.1007/s11709-013-0201-2

Zhang, Y., Yue, P., Zhang, G., Guan, T., Lv, M., Zhong, D., 2019. Augmented Reality Mapping of Rock Mass Discontinuities and Rockfall Susceptibility Based on Unmanned Aerial Vehicle Photogrammetry. Remote Sensing 11. https://doi.org/10.3390/rs11111311

Zhao, D., Jiao, Y., Wang, J., Ding, Y., Liu, Z., Liu, C., Qiu, Y., Zhang, J., Xu, Q., Wu, C., 2020. Comparative performance assessment of landslide susceptibility models with presence-only, presence-absence, and pseudo-absence data. J. Mt. Sci. 17, 2961–2981. https://doi.org/10.1007/s11629-020-6277-y

Zhou, C., Yin, K., Cao, Y., Ahmed, B., Li, Y., Catani, F., Pourghasemi, H.R., 2018. Landslide susceptibility modeling applying machine learning methods: A case study from Longju in the Three Gorges Reservoir area, China. Computers & Geosciences 112, 23–37. https://doi.org/10.1016/j.cageo.2017.11.019

Zhou, Z., Cai, X., Chen, L., Cao, W., Zhao, Y., Xiong, C., 2017. Influence of cyclic wetting and drying on physical and dynamic compressive properties of sandstone. Engineering Geology 220, 1–12. https://doi.org/10.1016/j.enggeo.2017.01.017

Zhu, A.-X., Miao, Y., Liu, J., Bai, S., Zeng, C., Ma, T., Hong, H., 2019. A similarity-based approach to sampling absence data for landslide susceptibility mapping using data-driven methods. CATENA 183, 104188. https://doi.org/10.1016/j.catena.2019.104188

Zhu, A.-X., Miao, Y., Yang, L., Bai, S., Liu, J., Hong, H., 2018. Comparison of the presence-only method and presence-absence method in landslide susceptibility mapping. CATENA 171, 222–233. https://doi.org/10.1016/j.catena.2018.07.012

Zimmer, V.L., Collins, B.D., Stock, G.M., Sitar, N., 2012. Rock fall dynamics and deposition: an integrated analysis of the 2009 Ahwiyah Point rock fall, Yosemite National Park, USA. Earth Surface Processes and Landforms 37, 680–691. https://doi.org/10.1002/esp.3206