



# A novel dynamic rockfall susceptibility model including precipitation, temperature and snowmelt predictors: a case study in Aosta Valley (northern Italy)

**Abstract** The overarching goal of the study was the development of a potentially dynamic rockfall susceptibility model by including climate predictors. The work is based on previously defined critical thresholds relating three climate indices — effective water inputs (EWI), wet-dry cycles (WD) and freeze–thaw cycles (FT) — and rockfall occurrence. The pilot area is located in the Aosta Valley region (Italian Western Alps). The susceptibility model settings were optimized through a stepwise procedure, carried out by means of generalized additive models (GAM). Predictors included topographic, climatic and additional snow-related variables. As climatic predictors, the mean annual threshold exceedance frequency was calculated for each index. All models were developed including an automatic penalization of statistically non-significant variables (i.e. shrinkage). The initial susceptibility model was set without considering potential inventory bias. Secondly, a “visibility mask” was produced to limit the modelling domain according to the rockfall event census procedures. Thirdly, GAMs functional relationships were analysed to verify the physical plausibility of predictors. Finally, to reduce concurvity, a principal component analysis (PCA) including climatic and snow-related predictors was carried out. Key findings were as follows: (i) ignoring inventory bias led to excellent model performance but to physically implausible outputs; (ii) the selection of non-rockfall points inside a “visibility mask” is effective in managing inventory bias influence on outputs; (iii) the inclusion of climate predictors resulted in an improvement of the physical interpretability of the associated models and susceptibility maps, being EWI, WD and the maximum cumulated snow melting the most important physically plausible climate predictors; (iv) the PCA strategy can efficiently reduce model concurvity.

**Keywords** Alps · Threshold exceedance frequency · Climate variables · Inventory bias · Generalized additive models · Principal component analysis

## Introduction

Due to the high energy and mobility, rockfalls are a major cause of fatalities (Hoek 2000; Frattini et al. 2008) and deeply affect human society and infrastructures in mountainous environments (Raveland and Deline 2010; Duvillard et al. 2015; Scavia et al. 2020). In the lower portion of slopes or along valley bottoms, rockfalls may damage buildings, roads, rail routes and other properties. At higher altitudes, these events mainly 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 aware of the increasing rockfall-related risks (Magnin et al. 2017), also in consideration of the effects of the twenty-first century global warming (Gobiet et al. 2014; Stoffel et al. 2014). For this reason, deciphering the role of climatic factors in rockfall predisposition and triggering is a fundamental research challenge that needs to be addressed to elaborate adaptation strategies (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), this should be the main thread guiding studies on landslide susceptibility, hazard and risk. Susceptibility is defined as the as the likelihood of a mass movement (in this case a rockfall) occurring in an area based on its characteristics (Brabb 1984).

As found out in the review of Reichenbach et al. (2018), until 2016, the inclusion of climate-related variables in susceptibility models was quite rare, with only 2.8% of them including a precipitation-related predictor and only 0.3% including other climatic predictors. This may be linked to how susceptibility is generally perceived; indeed, the traditionally used geo-environmental predictors (among all, those derived from the DEM) are considered static and, consequently, susceptibility is inherently deemed as stationary (Lombardo et al. 2020). In addition, 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. However, the validity of the assumption that climate processes would not affect landslide spatial occurrence is suitable only over small areas, where the conditions are homogeneous, and not for large or complex areas, where different microclimatic conditions may occur (Catani et al. 2013). Once recognized that climate-related processes affecting slope instability are dynamic in time and space, synthetic process-related non-stationary variables, adaptable to reveal climate (and more specifically climate change) impacts, are necessary (Camera et al. 2021).

In the most recent years, researchers started addressing this gap focusing mainly on the introduction of rainfall-related variables in shallow landslide susceptibility models. As synthesized in Camera et al. (2021), three common approaches can be found in literature. The first approach consists in the inclusion of precipitation in susceptibility models in the form of average-related metrics (e.g. mean annual rainfall, mean monthly rainfall and rainy days frequency). The second approach involves the development of models based on specific rainfall events summarized as multiple-day maximum cumulated precipitation variables. The third approach combines

stationary susceptibility models with either precipitation thresholds exceedance or an additional temporal statistical model including antecedent cumulated rainfall and soil saturation degree. Only few researchers attempted to further synthesize 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 and duration. 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 snowmelt events occurring in a hydrological year. The role of temperature on stability — and more in general its effects on the hydro-mechanical behaviour of rock masses — is mainly explored through physically based models (e.g. Gunzburger et al. 2005; Grämiger et al. 2018; Morcioni et al. 2022). For data-driven models targeting landslide susceptibility, the only example including temperature explicitly is the study of Loche et al. (2022). The authors focused on the seismic region of Wenchuan, affected by an earthquake in 2008, and found that land surface temperature can explain post-seismic landslide patterns. However, research gaps remain about the inclusion in landslide susceptibility models of climatic predictors other than rainfall and the development of models for events other than shallow landslides (e.g. rockfalls).

Another open question in landslide susceptibility analysis relates to the physical and geomorphic plausibility of the optimized statistical model (Steger et al. 2016, 2021; Camera et al. 2021; Bajni et al. 2022) — i.e. to the coherence of predictor effects on landslide susceptibility values (model outputs) with respect to the available knowledge and observations about the physical processes triggering landslides. Physical and geomorphic plausibility is strictly linked to the characteristics of the landslide inventory available and of possible related bias and inaccuracy (Steger et al. 2021; Jacobs et al. 2020; Bajni et al. 2022). To overcome inventory limitations, some recent works focused on incorporating strategies in the modelling procedure to limit bias effects. 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. 2016, 2017). Other works are based on the development of a specific sampling strategy for non-landslide points, limiting the modelling domain to monitored areas (Knevels et al. 2020; Bornaetxea et al. 2018). Moreover, the consistency and coherency of predictors behaviour in the model with their physical role on landslide occurrence is fundamental 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. 2022). 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 of data-driven model outcomes is crucial in the selection of the best, most suitable model in a multitude of possible, equal-performing ones.

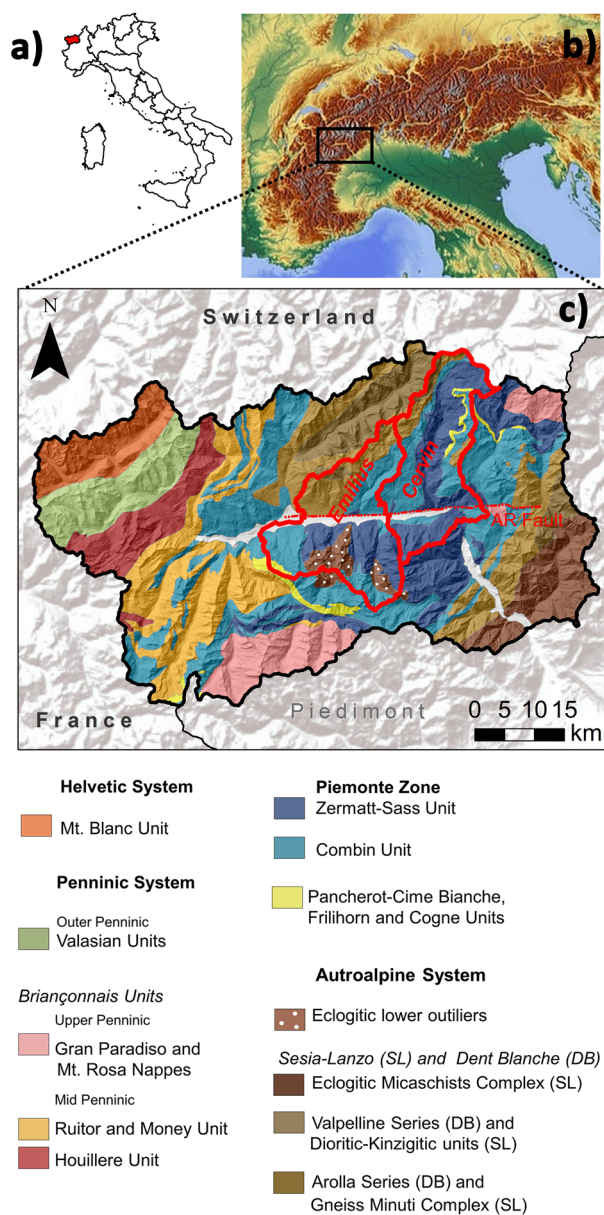
Finally, it is a common practice to carry out a multicollinearity analysis among predictors before using them as independent variables in a statistically based environmental models, both when

dealing with logistic regression and complex machine learning models (Camera et al. 2017; Chen 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 multicollinearity issues between each other and with morphometric predictors have to be expected. A possible way to overcome these issues may derive from dimensionality reduction techniques (e.g. principal component analysis, factor analysis, feature extraction), in order to obtain a set of uncorrelated predictors for environmental modelling and analysis (e.g. Sabokbar et al. 2014; Messenzehl et al. 2018; Mancini et al. 2019; Li et al. 2021; Zhu et al. 2022).

The overarching goal of the study was to test the influence of climate-related spatially distributed predictors on rockfall susceptibility in an Alpine environment. The study focused on 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. In a previous study (Bajni et al. 2021), the authors identified specific climatic processes relevant for rockfall occurrence in the area and summarized them in three indices, namely effective water inputs (including rainfall and snowmelt; EWI), wet-dry episodes (WD) and freeze–thaw cycles (FT). Starting from a weather station-based dataset, the authors calculated the statistical distribution of the three climate indices for the period 1990–2018 and defined, for each of them, empirical thresholds related to rockfall occurrence. Thresholds were designed on the concept of *non-ordinary* climatic conditions, defined as the 75<sup>th</sup>–90<sup>th</sup> percentile of the statistical distribution exceeded by at least 50% of the rockfall population analysed. The present study represents the prosecution of the work of Bajni et al. (2021) and has the following specific objectives: (i) to introduce potentially dynamic climatic predictors in a rockfall susceptibility model; (ii) to develop a sampling strategy able to deal with data collection and inventory bias issues within the model; (iii) to determine an effective procedure to model rockfall susceptibility able to guarantee the physical plausibility of the relationships between modelled predictor and susceptibility response by exploiting the interpretability and flexibility of generalized additive models (GAM) and iv) to limit potential multicollinearity effects between predictors exploring advantages and limitations of substituting the outputs of a principal components analysis to the original climatic predictors set.

### Study area

The Aosta Valley region is located in the North-western Italian Alps and represents one of the main alpine valley systems (Fig. 1a, b). The territory extends through the Europe-verging structural domains of the Western Alps (Fig. 1c), crossing the complex pile of continental nappes and minor ophiolitic leaves, characterized by ductile deformation and a differential subduction-related metamorphism, ranging from the blueschist to the eclogitic facies, and locally ultra-high pressure conditions (Frey et al. 1999; Bistacchi et al. 2001). From top to bottom, the pile consists of the Adria-derived Austroalpine system, the ophiolitic Piemonte zone and the Europe-derived continental nappes (Argand 1911; Dal Piaz 2001; Bistacchi et al. 2001; Giusti et al. 2004; Ballèvre et al. 2015; Ellero and Loprieno 2017; Tartarotti et al. 2019). Following the Alpine orogenesis, this complex overlapping structure was subject to neo-tectonic brittle



**Fig. 1** a Location of the Aosta Valley region in Italy and b in the Alps (image from <https://pixabay.com/illustrations/italy-alps-alpine-region-map-1804893/>). c Geo-structural map of Aosta Valley (modified from Bigi et al. 1990; Ellero and Loprieno 2017) and study area (red-bordered). AR fault, Aosta Ranzola Fault System

dislocation processes, attributable to the Aosta-Ranzola fault system (Ballèvre et al. 1986; Bistacchi et al. 2001).

The Aosta Valley topography plays an important role in defining the regional climate. Mountain ranges (bordering the North and South of the region) are characterized by average rainfall amounts above 1400 mm/y and winter snowy precipitation; their presence also hamper the access of Mediterranean and Atlantic humid air masses inducing a high degree of aridity in the central main E-W oriented valley axis (Orusa and Borgogno Mondino 2021). According to Rubel et al. (2017), despite the limited extension of the region,

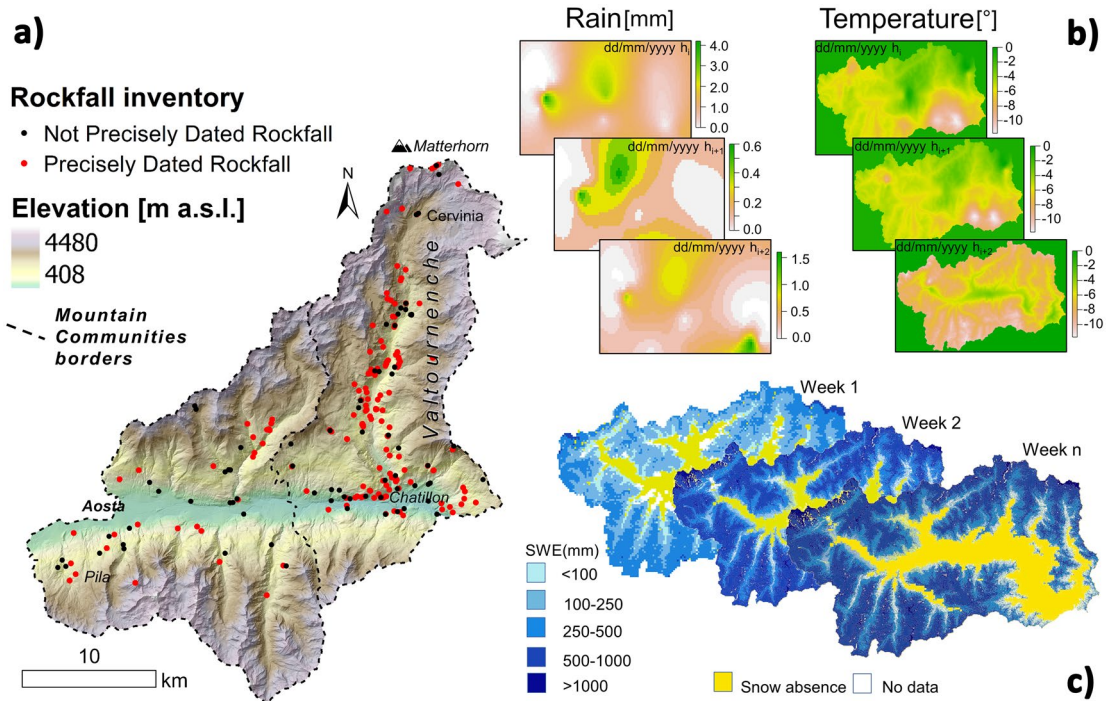
different climate regime coexists in the area, from alpine tundra to warm temperate until arid-semiarid.

The actual study area (Figs. 1c and 2a) is located in the central portion of the region and consists of the Mountain Communities of Mont Cervin and Mont Emilius, covering an area of 670 km<sup>2</sup> and an altitude range between 400 and 4500 m a.s.l. The Mont Cervin Mountain community is characterized by a main N-S oriented valley (named Valtournenche), which connects the Matterhorn hiking area and the highly frequented town of Cervinia (a popular ski area) in the North, to the main regional E-W oriented valley where the highway connecting Italy to France lies. The Mont Emilius Mountain Community extends from the southern massifs (where the Ski Area of Pila is located) to the northern mid-altitude slopes overlooking the main regional E-W oriented valley axis, from Chatillon to Aosta. The study area is characterized by the presence of different lithologies belonging prevalently to the ophiolitic Piemonte Zone and Austroalpine outliers. In the southern part of the study area, metaophiolite-related lithologies of the Zermatt-Saas zone (mainly serpentinites, gabbros, amphibolites, prasinites) are abundant together with eclogitic terms belonging to the Mont Emilius Austroalpine outlier (mainly micashists). In the northern part of the Mont Emilius Mountain Community, the meta-ophiolite lithologies of the Combin Zone (mainly calceschists, marbles, quartzites) are prevalent, together with the Mont Mary Austroalpine upper outlier (mainly ortogneisses and metagranitoids in green schist facies). The Valtournenche valley bottom is prevalently composed by Zermatt-Saas and Combin zone lithologies, whereas the northern valley head is characterized exclusively by the Dent Blanche and Mont Cervin Austroalpine upper outliers (mainly ortogneisses and metagranitoids in green schist facies too).

## Data

The rockfall inventory (Fig. 2a) used for this study was derived from Bajni et al. (2021). For the reference period (1990–2018), it consists of 243 rockfall records extracted from the publicly available Landslide Regional Database (<http://catastodissesti.partout.it>), for which the source areas were thoroughly checked and validated through a geo-morphic analysis (details in Bajni et al 2021). Rock-falls were identified as detachment scarp representative points and are associated with a unique ID. 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.

Regarding morphometric data, the digital terrain model (DTM) made available by Aosta Valley region (<https://geoportale.regione.vda.it/mappe/informazionigeoscientifiche>) was used as base. It has a 2 m × 2 m horizontal resolution but for this study, it was resampled at 10 m × 10 m, using ESRI ArcGIS<sup>®</sup> 10.2.2 Spatial Analyst tools. To derive geological information, the regional geological–geomorphological map at the 1:10,000 scale was used. The map is publicly consultable on the Aosta Valley geoportal WebGIS (<http://geologiavda.partout.it/cartaGeologicaRegionale?!=it>), while the associated shapefiles were made available upon request by the Regional Geological Office. The map reports different information levels, including the main geo-structural domains with a very detailed description based on lithologic characteristics and metamorphic imprint. Also, the map reports quaternary deposit types, including “large boulder deposits”,



**Fig. 2** **a** Available rockfall events locations in the selected study area (from Bajni et al. 2021). **b** Schematic visualization of the available hourly grid-based precipitation and temperature datasets (Pignone and Rebora 2014). **c** Schematic representation of the available grid-based SWE (snow water equivalent) dataset (Filippa et al. 2019)

and linear geological–geomorphological features, including some landslide “detachment scarps” (not dated).

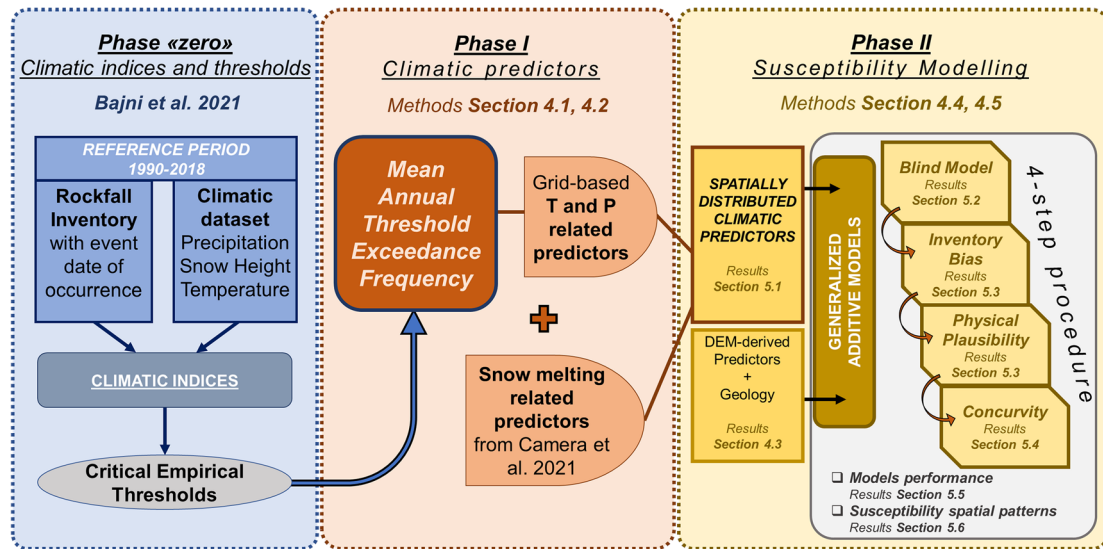
The Centro Funzionale Regione Autonoma Valle d’Aosta made available for this study a raster dataset (not published) of precipitation (mm) and temperature (°C) for the whole Aosta Valley region, for the period 2003–2020. The dataset has an hourly temporal resolution and a spatial resolution of  $0.00129^\circ$  (about 122 m). Each map is derived from the spatial interpolation of either precipitation data recorded at rain gauges (with no distinction between solid and liquid, thus including snow melting inputs for heated rain gauges) or temperature data acquired through thermometers installed at meteorological stations located on the regional territory. Stations location and data are accessible at [https://presidi2.regione.vda.it/str\\_dataview](https://presidi2.regione.vda.it/str_dataview). In particular, precipitation values were interpolated adopting the GRISO Rainfall generator (Pignone and Rebora 2014). An example of the available gridded datasets is provided in Fig. 2b. In addition, the Agenzia Regionale per la Protezione dell’Ambiente (ARPA Valle d’Aosta) 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 2020 (Filippa et al. 2019). The dataset can be viewed on the ARPA website (<https://www.arpa.vda.it/it/effetti-sul-territorio-dei-cambiamenti-climatici/neve/swe>) and was made available upon specific request. It is limited to the winter months (November–May) of each hydrological year, it 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 interpolation of the snow height factor, the dataset does not guarantee the conservation of mass between consecutive SWE maps (Camera et al. 2021). An example of the non-processed SWE datasets is provided in Fig. 2c.

### Methodology

The methodology followed three main phases (Fig. 3) and summarized as follows:

- (i) Phase *zero* is the propaedeutic previous work of Bajni et al. (2021), where the authors identified three climatic indices relevant for rockfall occurrence in the study area and defined empirical critical thresholds for each of them based on the concept of *non-ordinary* climatic conditions. The three climatic indices were as follows: (1) EWI, defined as the cumulated amount of effective water inputs (i.e. rainfall plus snowmelt); (2) WD, defined as wet-dry episodes, where a wet episode coincides with an hourly rainfall amount  $\geq 0.2$  mm and a dry episode occurs when in 24 consecutive hours no precipitation is detected; (3) FT, defined as freeze–thaw cycles, where a cycle is defined as a temperature transition across  $0^\circ\text{C}$ .
- (ii) The first phase concerned the formulation of synthetic, spatially distributed climatic predictors to be included in a rockfall susceptibility statistical model. To this purpose, the three climate-related indices and associated empirical criti-



**Fig. 3** Synthetic workflow and procedural steps adopted

cal thresholds relevant for rockfall occurrence as defined in Bajni et al. (2021) were converted into mean annual threshold exceedance frequencies, with a specific calculation for each climate index. Differently from the previous work of Bajni et al. (2021) in which meteorological data from the regional weather stations network were used, in this study, the raster dataset of precipitation and temperature (“Data” section) for the period 2003 to 2020 was preferred and assumed representative of the 30-year period climate normal ranging from 1991 to 2020 (WMO 1989, 2007). Moreover, two additional snow dynamics related predictors were available for the same study area from the previous work of Camera et al. (2021) and thus introduced in the model.

- (iii) The second phase concerned the implementation of a rockfall susceptibility model through generalized additive models (GAM) including climate predictors. Attention focused on the investigation of the climate predictors role, their 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 (Fig. 3), starting from a so-called blind model (i.e. a susceptibility model created without awareness of the rockfall inventory characteristics and of the physically driven processes potentially influencing susceptibility).

#### Calculation of climatic threshold exceedance frequency

To develop the climate-related predictors, an approach similar to Camera et al. (2021) was implemented. For each index, starting from a time series of either precipitation or temperature, the mean annual threshold exceedance frequency ( $TEF_a$ ) was calculated as:

$$TEF_a = \frac{n}{Ndays} * 365 \quad (1)$$

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 under analysis. The beginning of a rainfall event was set when a non-zero rainfall occurred (i.e. a minimum of 0.2 mm), while its end was imposed after 24 h without precipitation. 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 based on an intensity–duration threshold normalized on rainy days, using the available hourly precipitation dataset aggregated at a daily timestep. From the aggregated daily precipitation time series, the 1-, 3-, 7-, 15-, 30- and 60-day antecedent cumulated precipitation was calculated. Then, for each day, it was verified if any of the 1- to 60-day cumulated precipitation exceeded the threshold defined for the index (Bajni et al. 2021). Finally, each day was attributed a value of either “1” or “0” for threshold exceedance and non-exceedance, respectively. To ensure the independence between consecutive exceedance events, at least 24 h of non-exceedance (i.e. “0” values) must be recorded between successive “1” values; otherwise, they were counted as a single exceedance event. Following, the parameter  $n$  in Eq. 1 for the EWI index was obtained for each pixel by summing all the threshold exceedance events (i.e. the values equal to one in the data series) in the period from 2003 to 2020.

The  $TEF_a$  for the WD index was calculated on the grid-based hourly precipitation dataset. The threshold selected for the calculation was the “global threshold” (i.e. the one derived from non-ordinary conditions established from a statistical distribution derived using data from all the available weather stations) defined in Bajni et al. (2021). For each timestep (i.e. 1 h) in the time series,

the numbers of WD episodes recorded in the antecedent 30-, 60-, 90-, 120-, 180- and 365-day durations were calculated. Secondly, it was verified if, for the considered time step, the number of episodes recorded in the antecedent selected time frames (i.e. 30 to 365 days) exceeded the threshold defined for the WD index (Bajni et al. 2021) for at least one of the considered durations. Finally, each time step was attributed a value of either “1” or “0” for threshold exceedance or non-exceedance, respectively. As for the EW index, the parameter  $n$  in Eq. (1) was obtained summing the exceedances in the period from 2003 to 2020.

The  $TEF_a$  for the FT index was calculated on the grid-based hourly temperature dataset. For coherency with the other indices, the calculation was based on an intensity–duration like threshold (i.e. intensity intended as the number of freeze–thaw cycles occurring in predefined duration time periods) normalized on FTN (freeze–thaw normal, Bajni et al. 2021). For every hour, the numbers of FT episodes recorded in antecedent 1-, 3-, 7-, 15-, 30-, 60-, 90-, 120-, 180- and 365-day periods were calculated. Differently from the other indices, to ensure the independency of the threshold exceedance events, each duration was treated separately and then aggregated as detailed below.

For the 1-day duration, the procedure was as follows. The numbers of cycles accumulated in the 24 h preceding the considered timestep were calculated and each timestep was attributed a value of either “1” or “0” if the number of cycles exceeded or not the threshold defined for the index. Subsequently, a threshold exceedance event was counted only if the “1” values were temporally distant at least 24 h (i.e. 1 day). For 3-day to 365-day durations, a “cascade” incremental procedure was implemented. This means that for a given duration (e.g. 3 days), a freeze–thaw cycle can contribute only to one exceedance for that duration time-period (e.g.  $\pm 72$  h from the time it has been recorded) and if the threshold at a specific time is exceeded for multiple durations (e.g. 1-day and 3-day duration), it still counts “1”. Finally, the parameter  $n$  in Eq. (1) for the FT index was obtained summing all the independently counted threshold exceedance events. Additional details on the calculation procedure are given in the Supplementary material.

### Snow melting predictors

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 presented in the “Data” section. 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. Following, the consistency between the snow dynamics reproduced by the raster dataset and the continuous data recorded at meteorological stations spread over the study area was verified by means of correlations (for additional details, see Camera et al. 2021).

### Other geo-environmental predictors

Traditional DEM-derived predictors (van Westen et al. 2008; Reichenbach et al. 2018) were included in the rockfall susceptibility model. They were elevation, slope, aspect — included as northness =  $\cos(\text{aspect})$  and eastness =  $\sin(\text{aspect})$  — profile curvature, plan curvature and SAGA Topographic Wetness Index (SWI). They were derived at a 10 m  $\times$  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 (1:10,000 scale) available for the study area (refer to “Data” section). The distinction adopted followed a combined lithological and metamorphic criterion, as the widely variable metamorphic imprint characterizing the study area may differentiate the intact rock strength and rock mass properties. The classes were as follows: (i) Zermatt-Saas ophiolites in the eclogitic facies, (ii) Austroalpine upper outliers in the blueschist facies, (iii) Combin ophiolites in the blueschist facies, (iv) Austroalpine lower outliers in the eclogitic facies. The class adopted as reference for modelling purposes was the most abundant in the study area (i.e. the Zermatt Saas ophiolites).

### Rockfall susceptibility model setup and assessment

The 243 rockfall records available for the period 1990–2018 were used as the rockfall “presence” input of the binary response variable for the susceptibility modelling phase. The availability of the date of occurrence (or period of occurrence) within the rockfall inventory was a crucial characteristic for the calculation of EW index, FT and WD climate indices (in Bajni et al. 2021) and therefore for the inclusion of climate-related predictors in the susceptibility model in this study. Rockfall absence locations were selected by random extraction inside an “eligible” area, which is a subset of the modelling domain and varies according to the modelling step. The modelling domain was obtained from the study area excluding urban areas, glaciers, water bodies and quaternary deposits (i.e. areas where rockfalls cannot occur) as mapped in the geological-geomorphological map at the 1:10,000 scale. In step 1 (blind model), the “eligible area” was derived excluding from the modelling domain areas within a 100-m buffer from rockfall representative initiation points because considered probable rockfall source areas. From the geological–geomorphological map available for the modelling domain, 59 rockfall scarps were mapped. The analysis of the linear extension of these features returned mean and median values of 138 m and 108 m, respectively. Based on these data, the selected buffer dimension is considered representative of the rockfall source areas in the modelling domain. From step 2, the “eligible area” was further reduced through a visibility approach (details in the “Four steps to increase the physical plausibility of the statistical model” section). A 1:1 ratio between rockfall source areas “presence” and rockfall source areas “absence” points was adopted, as the percentage of the “eligible area” on the modelling domain was 97.6%. Hong et al. (2019) found this ratio optimal when the sampling area for absence data (i.e. “eligible” area) is quite large (i.e. the 99% of the modelling domain).

The rockfall susceptibility modelling was performed using generalized additive models (GAMs) through the R package *mgcv*

(Wood 2017). Variable selection was performed through the *shrinkage* option, which allows penalization of statistically non-significant predictors to the zero function and thereby selected them out of the model (Wood 2017). Predictors' behaviour was analysed through component smoothing functions (CSFs, i.e. the set of functions relating each predictor values to susceptibility values; Wood 2017) and odds ratios for geology categories. Both a non-spatial (nsCV) and a spatial (sCV) repeated *k-fold* cross-validation were carried out to estimate the area under the receiver operating characteristic (AUROC) curve. The nsCV involves a random selection of points to be attributed to the different folds, while the sCV involves the selection of points for the different folds based on k-means clustering of coordinates (Brenning 2012b). The nsCV suggests how well the model work within the study area (or areas with similar ranges of predictor values), while the sCV indicates the potential for model transferability outside the study area (i.e. to areas where predictor values could be outside of the training area range). 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 lower than 1 (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.

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

A four-step modelling procedure is proposed, based on stepwise modifications of the model setup to effectively deal with the effects of inventory collection procedures and potential deriving bias, predictors behaviour physical plausibility and predictors multicollinearity (concurvity). For all steps, the modelling algorithm and the model evaluation procedure remained fixed as described in the previous “Rockfall susceptibility model setup and assessment” section. Moreover, all steps included the development of a preliminary topographic model (i.e. a susceptibility model based exclusively on the DEM-derived predictors) before proceeding with the actual inclusion of climate and snow-related predictors. In this way, statistically non-significant (i.e. penalized) topographic predictors were excluded from the model, to reduce variables to a parsimonious set. All the developed models along with their characteristics are presented in Table 1.

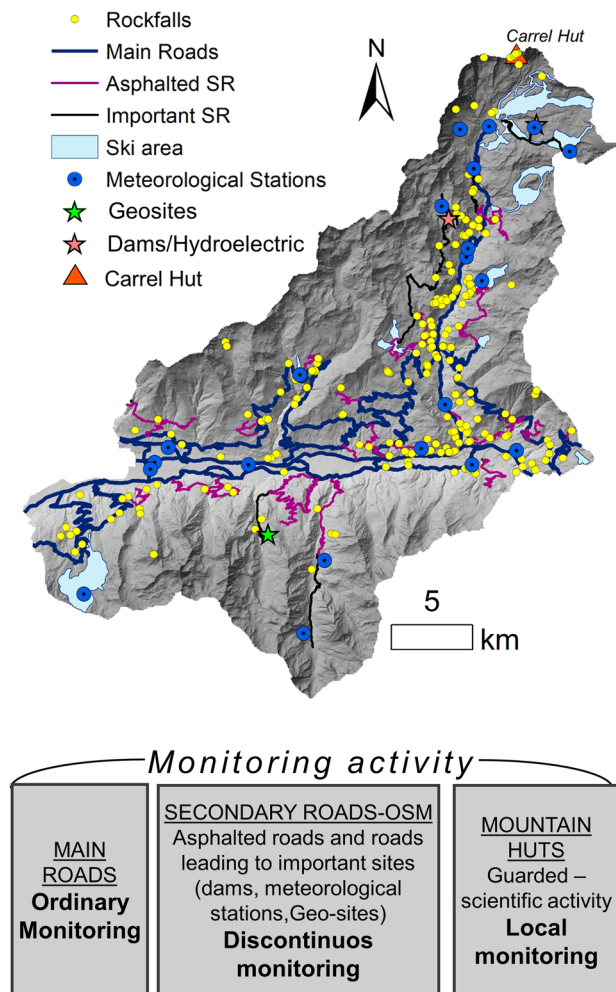
The first step of the modelling procedure consisted in a “Blind” model, i.e. a model that disregards a possible bias of the inventory and the physical plausibility of the relationships between predictor and susceptibility outcome. Only statistically non-significant (penalized) predictors were excluded. In detail, three “Blind” models with different predictors sets were produced, namely blind-TOPO, blind-CLIMATE and blind-GEO.

Step 2 was aimed at reducing the effects of the inventory bias on the model output reducing the “eligible area” for absence point selection by developing a visibility mask. As shown in Fig. 4, rockfalls are overrepresented close to road and infrastructure in comparison to remote areas of the study domain, most probably due to data collection procedures. Therefore, it was assumed that rockfall event reporting is dependent on the reachability of the sites. Principal

**Table 1** Model name, domain and predictors set used in the developed models

Model name	Modelling Domain	DEM-derived predictors	Climatic predictors	Geology	Step
<i>Blind-TOPO</i>	Eligible area	Elevation, slope, northness, eastness, SWI, profile curvature, plan curvature	-	-	1
<i>Blind-CLIMATE</i>	Eligible area	Non-penalized predictors from blind-TOPO	<i>EWI, WD, FT</i> <i>SWEep, SWEmax</i>	-	1
<i>Blind-GEO</i>	Eligible area	Non-penalized predictors from blind-TOPO	<i>EWI, WD, FT</i> <i>SWEep, SWEmax</i>	Reclassified geology	1
<i>Visibility-TOPO</i>	Eligible area + visibility mask	Elevation, slope, northness, eastness, SWI, profile curvature, plan curvature	-	-	2
<i>Visibility-CLIMATE</i>	Eligible area + visibility mask	Non-penalized predictors from visibility-TOPO	<i>EWI, WD, FT</i> <i>SWEep, SWEmax</i>	-	2
<i>Visibility-PP</i>	Eligible area + visibility mask	Non-penalized predictors from visibility-TOPO	<i>Non-penalized + physically plausible predictors from visibility-CLIMATE</i>	-	3
<i>Visibility-PCA</i>	Eligible area + visibility mask	Non-penalized predictors from visibility-TOPO	<i>PC1, PC2, ..., PCn</i>	-	4

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. Among secondary roads, only asphalted ones and those leading to point of interests (e.g. ski areas, dams and hydroelectric power plants, meteorological stations and geo-cultural sites) were retained. 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 road (either principal or retained secondary) was calculated together with the associated statistical distribution, represented in the form of a boxplot. The upper whisker of the boxplot (i.e. the largest value of the distribution that is no greater than the



**Fig. 4** Map showing principal and secondary roads and main infrastructures in the study area, together with the assumed scheme of the monitoring activities for new landslides (rockfalls) source area recognition and report

third quartile plus 1.5 times the interquartile range) was selected as the maximum “outer radius” parameter required by the visibility tool of ESRI ArcGIS® 10.2.2, here adopted to extract a visibility mask in which, reasonably, the monitoring designated personnel may record rockfalls visible along the driving routes (similarly to Bornaextea et al. 2018; Knevels et al. 2020). In addition to the obtained visibility mask, a buffer (buffer distance equal to the “Outer radius” selected for roads) centred in the Carrel hut (southern slope of Matterhorn) assumed frequently visited for technical-scientific purposes (presence of a micro-seismic monitoring system described in Amitrano et al. 2010; Occhiena et al. 2012) was added to the sampling scheme. The model was trained in the “eligible” areas, but the predictions are extended to the whole modelling domain. As for step 1, a preliminary parsimonious topographic model was developed (visibility-TOPO model) before proceeding with models including climate and snow-related predictors (visibility-CLIMATE model). For parsimonious, in general, it is meant a model with a limited number of variables explaining most of the variance; in this case, specifically, it means without non-significant penalized predictors.

Step 3 was aimed at optimizing the selection of predictors considering their physical behaviour by analysing their CSFs. Beside the quantitative performance, to keep a predictor within the model, it was considered essential that its effect on susceptibility had to be coherent with the known physical processes relating them to rockfall initiation. Therefore, an increasing susceptibility is expected for increasing threshold exceedance frequencies of EWI, WD and FT predictors. For snow-related predictors, an increasing susceptibility level is expected for high values of SWEep and high negative values of SWEmax (meaning high snowmelt). To obtain the physically plausible model (visibility-PP model), climate and snow predictors showing CSFs with non-plausible behaviours within the visibility-CLIMATE model were excluded.

Step 4 was aimed at reducing the concurrency among climatic predictors and producing a parsimonious model. According to Wood (2017), each smooth term in a model can be decomposed into two parts, one that lies in the space of one or more other terms in the model, and another part that is completely within the term’s own space. Concurrency varies from 0, indicating no shared space with other terms, and 1, indicating a total lack of identifiability. Concurrency is generally considered problematic for values above 0.8 (e.g. Camera et al. 2021). A principal component analysis was carried out, including the climatic and 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. Then, a PCA-based model (i.e. PCA-visibility) was produced by introducing in the model, as predictors, the first principal components explaining at least 70% of the total variance.

Finally, models were compared and discussed in terms of performance, climatic predictors’ behaviour and role, and 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”). Moreover, three susceptibility difference maps among steps were prepared to further discuss and analyse the different susceptibility spatial patterns.



## Results

### Synthetic climatic predictors

The developed climate and snowmelt-related predictors are presented in Fig. 5.

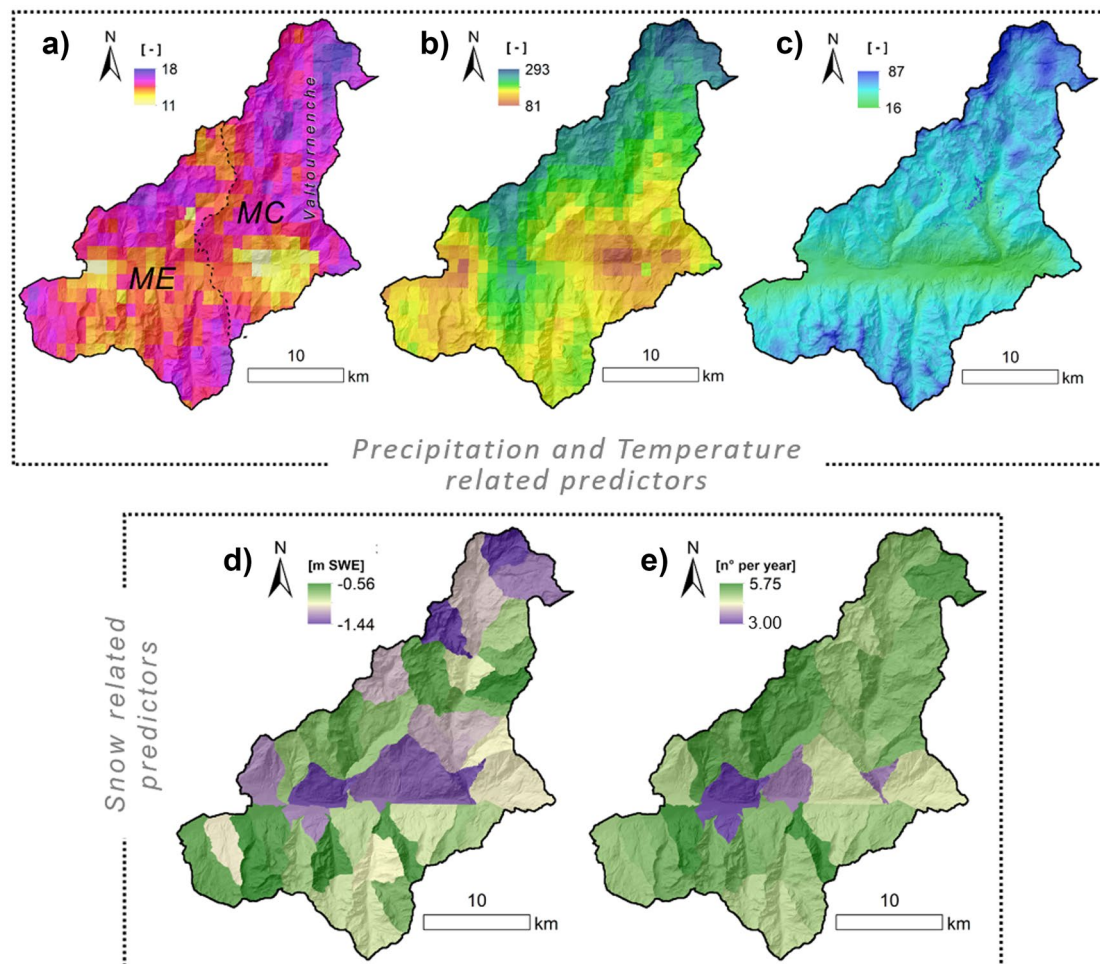
Figure 5a shows the spatial distribution of the EWI predictor, which is characterized by the highest values concentrated along the Valtournenche N-S axis, along the NW border of the study area and in its southern part. Figure 5b shows the spatial pattern associated to the WD predictor, which is characterized by the highest values in correspondence of the head of the Valtournenche Valley and by the lowest values in correspondence of the main E-W valley axis. The map also identifies a belt of high values in the mid-southern slope of Mount Emilius Mountain Community and along the NW border of the study area. Figure 5c shows the spatial pattern of the FT predictor, characterized by a substantial correlation with altitude. However, it was possible to observe some bands at mid-high altitudes where the threshold exceedance frequency showed very high values, close to the maxima; these zones may capture patterns

attributable to temperature inversion phenomena and the seasonal shifting of the zero isotherm.

The maps regarding the snowmelt predictors produced by Camera et al. (2021), namely SWEmax and SWEep, are presented in Fig. 5d and e, respectively. Note that for SWEmax, since snowmelt is a depletion of the SWE storage, negative values indicate snowmelt occurrence while positive values indicate snow accumulation.

### 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. 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. Penalization frequency and mDD% of each predictor in the blind models are collected in Table 2.



**Fig. 5** Climatic predictors maps: **a** EWI, effective water inputs. The dotted line represents the border between Mont Emilius Mountain Community (ME) and Mont Cervin Mountain Community (MC). **b** WD, wet and dry cycles and **c** FT, freeze thaw cycles. Snow melting predictors (Camera et al. 2021): **d** SWEmax, i.e. the maximum amount of melting recorded over 32-day periods in the whole data series and **e** SWEep, i.e. the average number of melting events occurring over 16-day periods in a hydrological year. ME, Mont Emilius Mountain Community; MC, Mont Cervin Mountain Community

In the blind-CLIMATE model (Fig. 6), the variable elevation remained the dominant predictor, followed by slope, EWl and profile curvature as the most important non-penalized variables (Table 2). The CSF for elevation, describing an almost linear and indirect relationship between elevation and rockfall occurrence, indicated the highest rockfall likelihood in correspondence of the lowest altitudes (approximately < 1800 m a.s.l.). This may be interpreted as a possible symptom of rockfall overrepresentation close to roads and principal infrastructures, thus attributable to a data collection effect. 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 is characterized by a slightly increasing CSF, indicating east-facing slope as mildly more susceptible than the west-facing ones. Regarding the CSF related to the climatic predictors (Fig. 6), the only predictor with a curve describing a physically plausible behaviour (and consistently maintained through the nsCV and sCV runs) was SWE<sub>max</sub>, indicating that the modelled landslide likelihood was highest in correspondence of areas with the highest cumulated snowmelt. The bell shapes of the CSF of EWl, WD and SWE<sub>ep</sub> are considered physically incoherent (Fig. 6). The FT predictor was frequently penalized in the CV runs, probably due to high concavity effects with elevation.

Accounting for the influence of geology (model blind-GEO), CSF values, predictors importance and CSF shape remained consistent. However, none of the odds ratios calculated had statistical significance (i.e.  $p$ -value > 0.05), which, in addition to a very low mDD% of 0.9%, denotes that the implemented geological predictor has very little influence on rockfalls in the study area. A possible cause of this model outcome may be the bias in the inventory, potentially leading to an underrepresentation of rockfalls in specific lithologies (e.g. the Austroalpine upper outliers, outcropping almost exclusively at high altitudes in the north-western part of the study area).

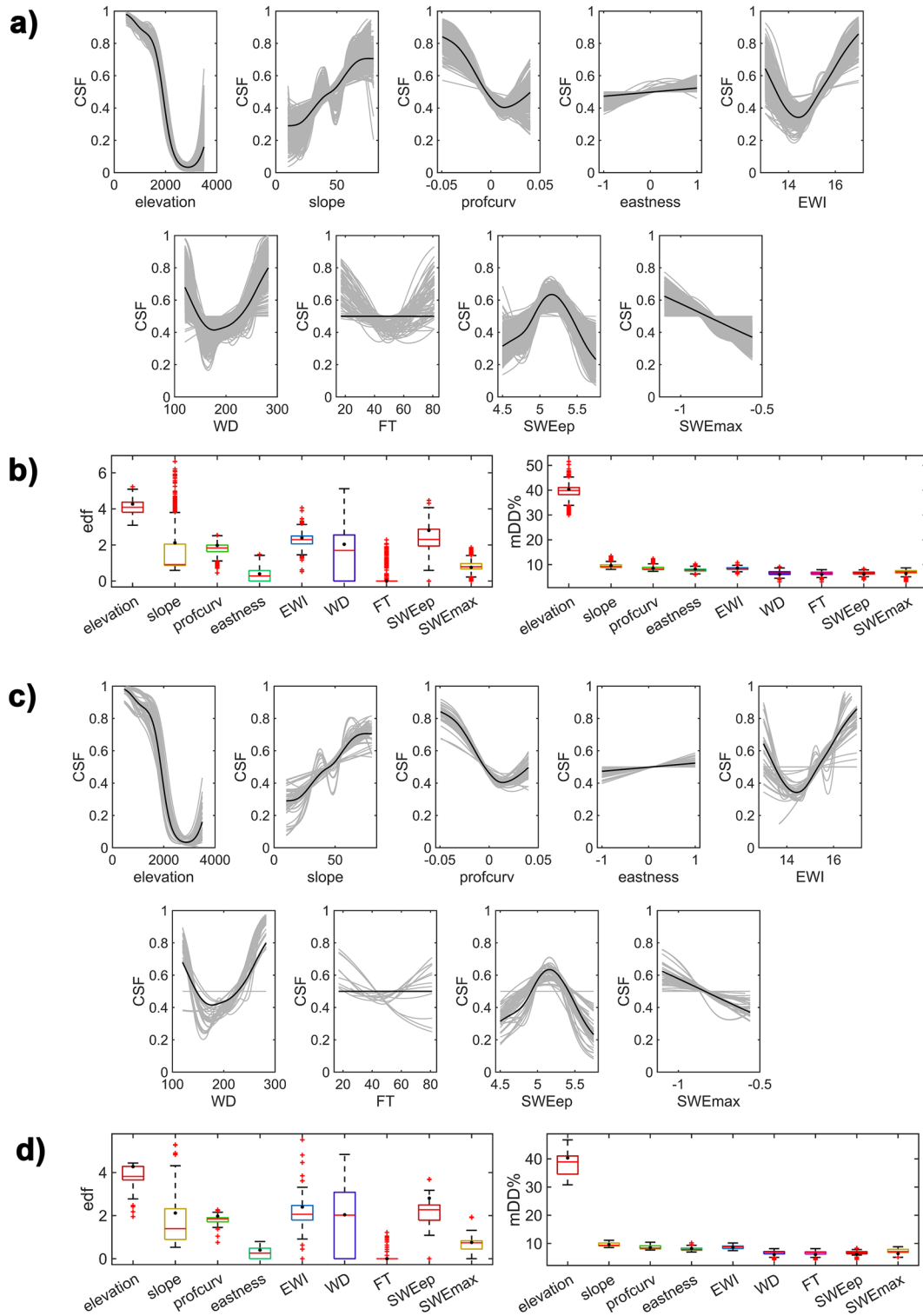
### The visibility mask and physically plausible models — step 2 and step 3

The boxplot associated to the statistical distribution of the distance between each rockfall and the closest road (either principal or retained secondary) is shown in Fig. 7a. The upper whisker of the boxplot selected as the “outer radius” parameter resulted equal to 871 m. The distance of the buffer centred in the Carrel hut location, in addition to the visibility mask produced from roads, was of 761 m and comprehended a total of 9 rockfalls. The final visibility mask, i.e. the effective modelling domain adopted from step 2 on, is shown in Fig. 7b. 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 rockfalls excluded were used as a holdout independent test set to additionally estimate model performance besides the cross-validation (“Rockfall susceptibility model setup and assessment” section). The remaining 201 rockfalls were associated to an equal number of non-rockfalls points, randomly selected inside the visibility mask with a 1:1 ratio. It is important to outline that it was not possible to include geology from step 2 on: by further limiting the study area with the aim of bias reduction, two geological 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, which hampers the possibility to transfer the trained model from the training to a test area (Guzzetti et al. 2006).

The topography-based model resulting from the application of the visibility approach (i.e. visibility-TOPO model) resulted in the penalization and consequent exclusion of the predictors eastness, plan curvature and SWl. To produce parsimonious models, these predictors were excluded from the visibility-CLIMATE and

**Table 2** Predictors’ penalization frequency and mDD% for the “Blind” models

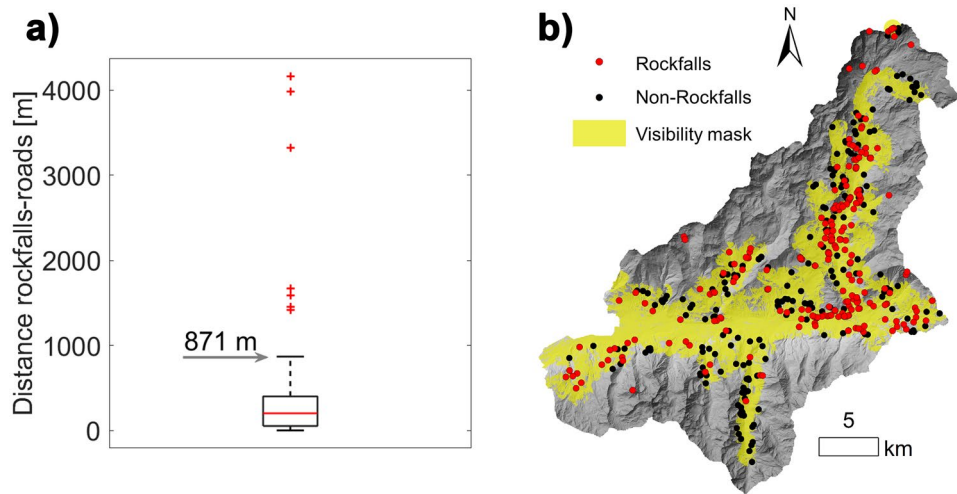
Model ID	Predictors	Penalization frequency (edf < 0.7) [%]		mDD%
		nsCV	sCV	
<b>Blind-TOPO</b>	<i>Elevation</i>	0.0	0.0	76.6
	<i>Slope</i>	3.4	7.2	7.2
	<i>Profile curvature</i>	0.2	0.0	5.9
	<i>Plan curvature</i>	97.6	91.2	4.6
	<i>Northness</i>	99.0	89.0	4.6
	<i>Eastness</i>	79.4	81.0	0.7
	<i>SWl</i>	86.6	80.0	0.4
<b>Blind-CLIMATE</b>	<i>Elevation</i>	0.0	0.0	40.3
	<i>Slope</i>	1.2	7.6	9.6
	<i>Profile curvature</i>	0.2	0.0	8.5
	<i>Eastness</i>	86.0	88.4	8.2
	<i>EWlind</i>	0.4	7.8	8.6
	<i>WD</i>	36.2	40.8	6.1
	<i>FT</i>	88.2	89.2	6.1
	<i>SWE<sub>ep</sub></i>	2.8	7.6	6.1
	<i>SWE<sub>max</sub></i>	32.6	37.2	6.4
<b>Blind-GEO</b>	<i>Elevation</i>	0.0	0.0	38.5
	<i>Slope</i>	0.6	4.2	10.3
	<i>Profile curvature</i>	0.0	0.0	9.1
	<i>Eastness</i>	91.4	85.4	8.1
	<i>EWlind</i>	2.2	7.4	8.5
	<i>WD</i>	10.2	20.4	5.9
	<i>FT</i>	85.4	84.8	5.9
	<i>SWE<sub>ep</sub></i>	4.0	13.8	6.8
	<i>SWE<sub>max</sub></i>	76.8	93.8	5.9
	<i>Geology</i>	-	-	0.9



**Fig. 6** Blind-CLIMATE model predictors component smoothing function (CSF), effective degrees of freedom (edf) and mean decrease in deviance explained (mDD%) for non-spatial CV (**a, b**) and spatial CV (**c, d**)

visibility-PP models. The penalization frequency, mDD% and concavity values of the models produced in step 2 and step 3 are summarized in Table 3.

As a confirmation of the robustness of the topographic predictors role across models, their CSF shapes (Fig. 8), penalization frequency and deviance explained remain almost unvaried and



**Fig. 7** a) Boxplot representing the distribution of the rockfalls distances from principal and secondary roads. b) Visibility masks and selected non-rockfall points

consistent in the three models (visibility-TOPO, visibility-CLIMATE and visibility-PP). The CSF of elevation shows 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. This confirms that the elevation-associated CSF in the blind models should be imputed to the bias in the inventory. However, even in the step 2 associated models, the low-altitude peak of the function is the highest, which 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. The CSFs for slope and profile curvature indicate that the modelled rockfall likelihood is highest for high inclined (approximately  $> 50^\circ$ ) and convex slopes. The CSF of northness is represented by a linear, gently inclined function (in some models very close to penalization) indicating that the modelled rockfall likelihood is 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 FT predictor resulted to be penalized on both the entire dataset and in most of the CV runs.

In the visibility-CLIMATE model, EWI appears as the most important climatic predictor with an mDD%  $> 10\%$ , followed by SWEmax with an mDD% of 5%. Both of these climatic predictors show a physically plausible behaviour (Fig. 8). The CSF related to EWI shows a nonlinear behaviour. At low values (between 14 and 15 threshold exceedances per year), the CSF presents a near-sigmoidal shape, probably due to the high dispersion of the data. 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 showing CSF coefficients below 0.5 (i.e. indicating a negative contribution to rockfall susceptibility). The CSF of SWEmax is a monotonous, almost linear function, coherent with the variable physical meaning. Conversely, WD and FT are penalized (CSFs represented by a horizontal line), while SWEep shows a physically implausible

behaviour (i.e. a double bell shape with two peaks); thus, it was excluded from the visibility-PP model. In the visibility-PP model (Fig. 8), despite the elimination of the physically implausible predictor SWEep, the CSF shapes, penalization frequency and deviance explained of EWI and SWEmax remain consistent and almost unvaried respect to the above-discussed visibility-CLIMATE model. As such, the FT predictor is penalized on both the entire dataset and in most of the CV runs. In the visibility-PP model, differently from the previous visibility-CLIMATE model, the WD predictor is not penalized; it shows a physically plausible CSF, represented by a monotonous, almost linear function, coherent with the variable physical meaning (Fig. 8).

Despite the statistically and physically based selection of predictors in the visibility-PP model, concurrency between them (Table 3) resulted quite high, being a possible cause of unstable estimates (Wood 2017). An evident concurrency issue is recorded between elevation and the FT predictor. It was not unexpected, as temperature variations are strictly coupled with altitudinal variations. For EWI, WD and SWEmax, concurrency values are medium–high.

#### Visibility-PCA model — step 4

A further reduction of concurrency, at least among 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 selected as predictors for the susceptibility models. The first principal component PC1 explained the climatic predictors FT, EWI and WD with a direct relationship (Fig. 9), meaning that high and positive PC1 values corresponded to high threshold exceedance frequencies. The second principal component PC2 explained the snow-related predictors SWEmax and SWEep with a direct relationship. This means that high positive values of PC2 corresponded to a high number of snowmelt events (i.e. high values of SWEep), but low values of cumulated snowmelt (i.e. SWEmax).

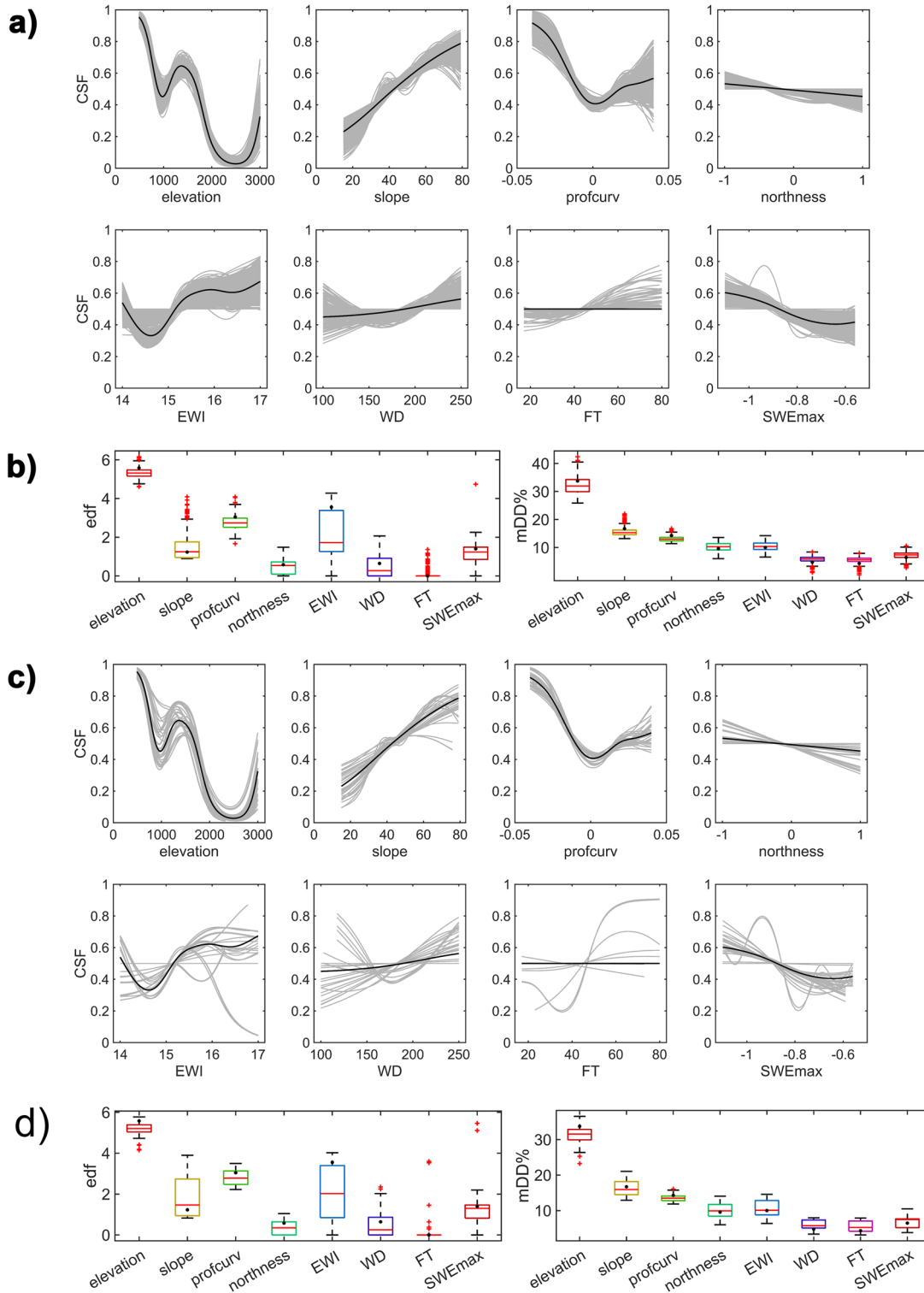
**Table 3** Predictors' penalization frequency, mDD% and concurrency for the "Visibility" models

Model	Predictor	Penalization Frequency (edf < 0.7) [%]		mDD%	Physical plausibility	Concurrency
		nsCV	sCV			
<i>Visibility-TOPO</i>	<i>Elevation</i>	0.0	0.0	42.0	<b>Yes</b>	0.24
	<i>Slope</i>	0.0	0.0	16.7	<b>Yes</b>	0.38
	<i>Profile curvature</i>	12.4	26.6	15.0	<b>Yes</b>	0.26
	<i>Plan curvature</i>	99.6	100.0	3.7	-	0.39
	<i>Northness</i>	0.0	0.0	9.3	<b>Yes</b>	0.25
	<i>Eastness</i>	94.4	89.2	6.8	-	0.23
	<i>SWI</i>	97.6	97.6	6.4	-	0.35
<i>Visibility-CLIMATE</i>	<i>Elevation</i>	0.0	0.0	28.7	<b>Yes</b>	0.90
	<i>Slope</i>	0.0	13.0	13.8	<b>Yes</b>	0.26
	<i>Profile curvature</i>	0.0	0.0	12.4	<b>Yes</b>	0.18
	<i>Northness</i>	57.8	78.8	10.3	<b>Yes</b>	0.48
	<i>EWI</i>	9.6	21.0	10.2	<b>Yes</b>	0.74
	<i>WD</i>	92.8	71.6	6.6	-	0.83
	<i>FT</i>	99.2	95.0	6.4	-	0.87
	<i>SWEep</i>	18.6	20.6	6.4	<b>No</b>	0.75
	<i>SWEmax</i>	9.0	11.4	5.1	<b>Yes</b>	0.76
<i>Visibility-PP</i>	<i>Elevation</i>	0.0	0.0	33.7	<b>Yes</b>	0.88
	<i>Slope</i>	0.0	0.0	16.7	<b>Yes</b>	0.22
	<i>Profile curvature</i>	0.0	0.0	14.3	<b>Yes</b>	0.15
	<i>Northness</i>	71.4	79.0	9.6	<b>Yes</b>	0.41
	<i>EWI</i>	8.4	21.0	10.0	<b>Yes</b>	0.70
	<i>WD</i>	67.0	58.2	4.8	<b>Yes</b>	0.77
	<i>FT</i>	97.8	95.2	4.3	-	0.85
	<i>SWEmax</i>	10.0	20.0	6.5	<b>Yes</b>	0.52

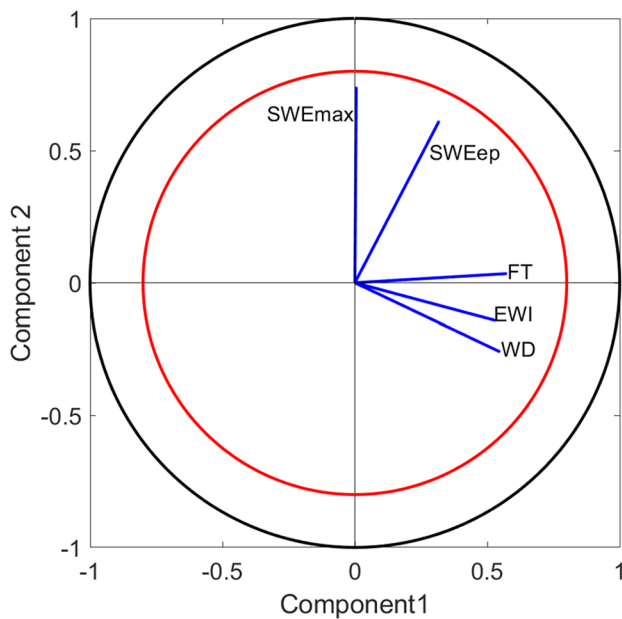
For the visibility-PCA model, the importance and CSFs of morphometric predictors (elevation, slope, profile curvature and northness) are consistently similar to the previously discussed models ("The visibility mask and physically plausible models — step 2 and step 3" section). Regarding the principal components, they show the expected behaviour (Fig. 10). The CSF of PC1 shows a near-sigmoidal shape for values below zero, probably inherited from the EWI index, while it shows an almost linear increasing behaviour for positive values. The 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, the effect of this predictor on the definition of PC2-related CSF is expected to be stronger than SWEep. As expected, the CSF of PC2 shows a linear monotonic decreasing function, physically coherent with SWEmax, which has 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 are reported in Table 4. In comparison to models from step 2 and step 3 (Table 3), concurrency between principal components is much lower than concurrency between climatic predictors. Also, calculating principal components allowed to reduce the concurrency of elevation with other predictors below 0.8.



**Fig. 8** Visibility-PP model predictors component smoothing function (CSF), effective degrees of freedom (edf) and mean decrease in deviance explained (mDD%) for non-spatial CV (**a, b**) and spatial CV (**c, d**). CSF shapes are valid for also for visibility-TOPO (elevation, slope, profcurv, northness) and visibility-CLIMATE (elevation, slope, profcurv, northness, EWI, FT, SWEmax)



**Fig. 9** 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 representation purposes, the components and variables scores were normalized between  $-1$  and  $1$ : the inner red circle represents a relationship of  $\pm 0.8$  and the outer black circle a relationship of  $\pm 1$

### Model performance

The variability of AUROC estimates from nsCV and sCV for all the tested model is presented in Fig. 11. The AUROCs indicated a similar and “excellent” (according to Hosmer et al. 2013 guidelines) discrimination capability for all the three “Blind” models, with a median AUROC scores on the test set between 0.8 and 0.9, both for the nsCV and the sCV. The lower performance of the blind-GEO model in comparison to the blind-climate model in the sCV confirms that the geological predictor reduced the discrimination capacity and spatial transferability of the model. Similar quantitative performances between training and test runs may induce to trust the “Blind” models. However, it has been demonstrated that this type of models may rely upon a physically non-plausible behaviour of predictors and on the influence of data collection procedures (i.e. altitude and, indirectly, vicinity to roads and infrastructures).

The results related to the nsCV showed “good” discrimination capability (Hosmer et al. 2013) for the three visibility-related models and the PCA-based model, with mean AUROCs between 0.75 and 0.80. All the models containing climate and snow-related predictors showed a slightly higher performance than the visibility-TOPO model. In terms of sCV, the discrimination capability of these models became only “acceptable” (Hosmer et al. 2013), with mean AUROCs between 0.6 and 0.7 for all models; the best performing model was the visibility-PCA model. Using the 42 rock-falls excluded from the visibility mask as an independent holdout

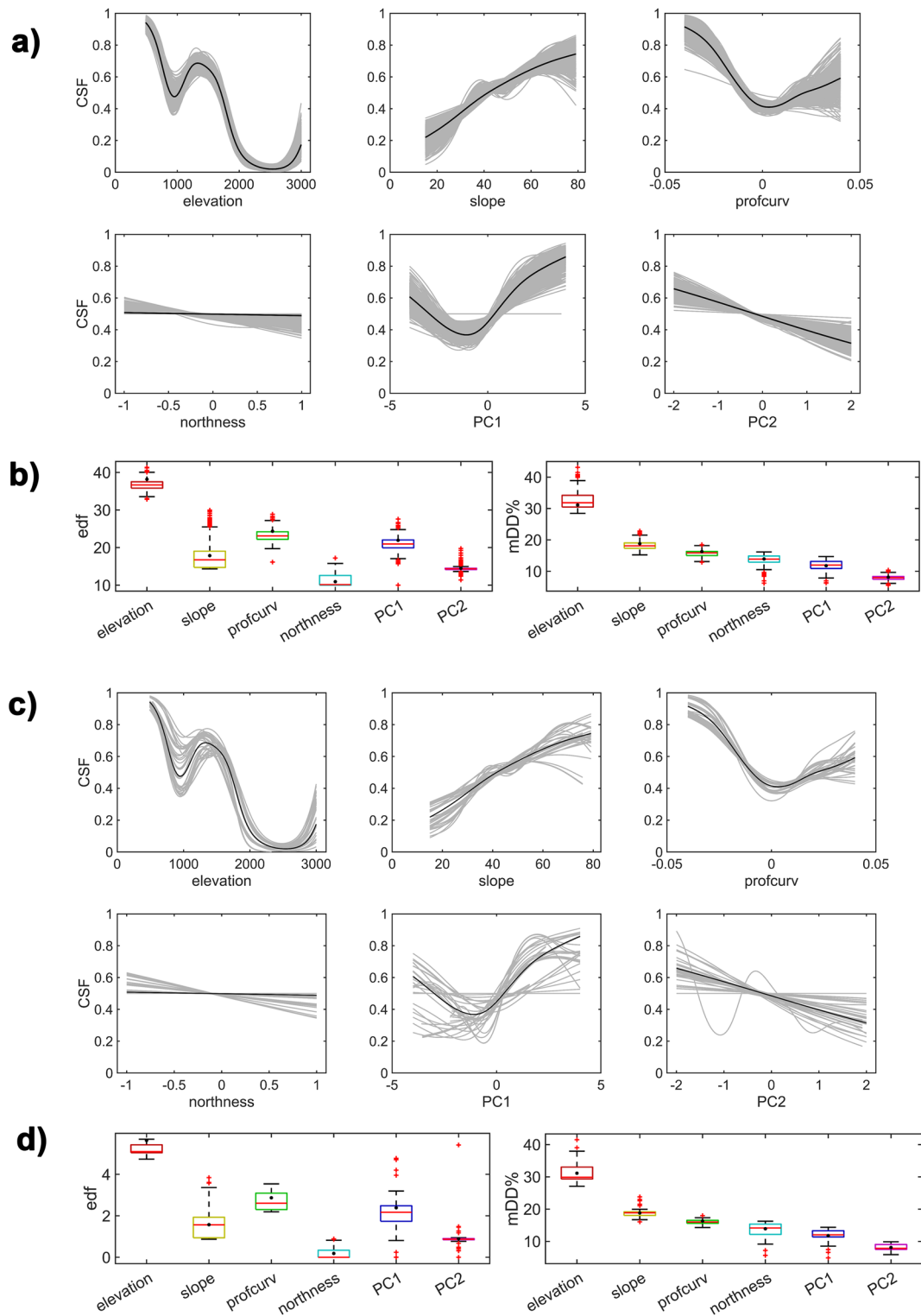
test set, the AUROC was 0.82 and 0.81 for the visibility-PP and the visibility-PCA model, respectively.

### Model comparison

A comparison of the spatial prediction patterns between the blind-CLIMATE, the visibility-PP and visibility-PCA derived susceptibility maps is provided in Fig. 12. These model outputs are shown with the intent to compare the spatial pattern of three potential “final” susceptibility maps (final maps intended as including all available predictor types, namely topographic and climatic predictors). The spatial pattern of the blind-CLIMATE model (Fig. 12a) reflects the behaviour of elevation as the bias-describing predictor. In detail, the susceptibility values are highest at the valley bottoms, close to roads and infrastructures, and show a gradual decrease with increasing elevation. This indicates that the susceptibility spatial pattern reflects the rockfall events collection procedures, rather than the environmental processes causing them. The susceptibility spatial pattern resulting from the visibility-PP model (Fig. 12b) benefits from the procedure adopted to reduce the inventory bias effects. A gradual decrease of susceptibility with elevation is not observed anymore, and the high susceptibility values still visible at low medium elevation can be interpreted as an actual presence of active slopes. Moreover, the map indicates medium to high susceptibility values in correspondence of high peaks and rock walls located in the northern and western part of the study area (head of the Valtournenche Valley, above 2000 m a.s.l.). In addition, the model records medium-high susceptibility values in lateral valleys located along the southern slope of the main E-W valley axis. The PCA-based model (Fig. 12c) reflects the same patterns observed in the visibility-PP susceptibility map.

Figure 13 has the double intent of pointing out the contribution of climate-related predictors in comparison to a model including only “common” topographic predictors (Fig. 13a) and of highlighting similarities and dissimilarities of “final” maps obtained with the step-wise model advancements (Fig. 13b, c). In detail, Map VAR-1 (Fig. 13a) reports the difference between the models visibility-PP and visibility-TOPO, thus highlighting the role of climatic predictors in the susceptibility spatial pattern. Positive values (i.e. higher susceptibility for visibility-PP than visibility-TOPO) can be noticed in almost the entire Valtournenche valley, which is characterized by the highest number of exceedances of both EWI and WD thresholds. Positive values are also present in some portions of the northern slopes facing the E-W main valley axis, which are characterized by the highest cumulated snowmelt (high negative values of SWEmax).

Map VAR-2 (Fig. 13b) reports the difference between the models visibility-PP and visibility-PCA; thus, it is linked to the different form in which climatic and snow-related predictors are included in the models. On the one hand, the variability in the spatial patterns can be explained by the absence of the SWEep (implausible behaviour) and FT (penalized) predictors in the visibility-PP model, which conversely contributed to the definition of the principal components in the visibility-PCA model. On the other hand, principal components explained about the 74% of the total variability associated to the climatic and snow-melting predictors, thus potentially missing a minor part of the associated signal.



**Fig. 10** Visibility-PCA model predictors component smoothing function (CSF), effective degrees of freedom (edf) and mean decrease in deviance explained (mDD%) for non-spatial CV (**a, b**) and spatial CV (**c, d**)



**Table 4** Predictors' penalization frequency, mDD% and concurvity for the PCA-based model

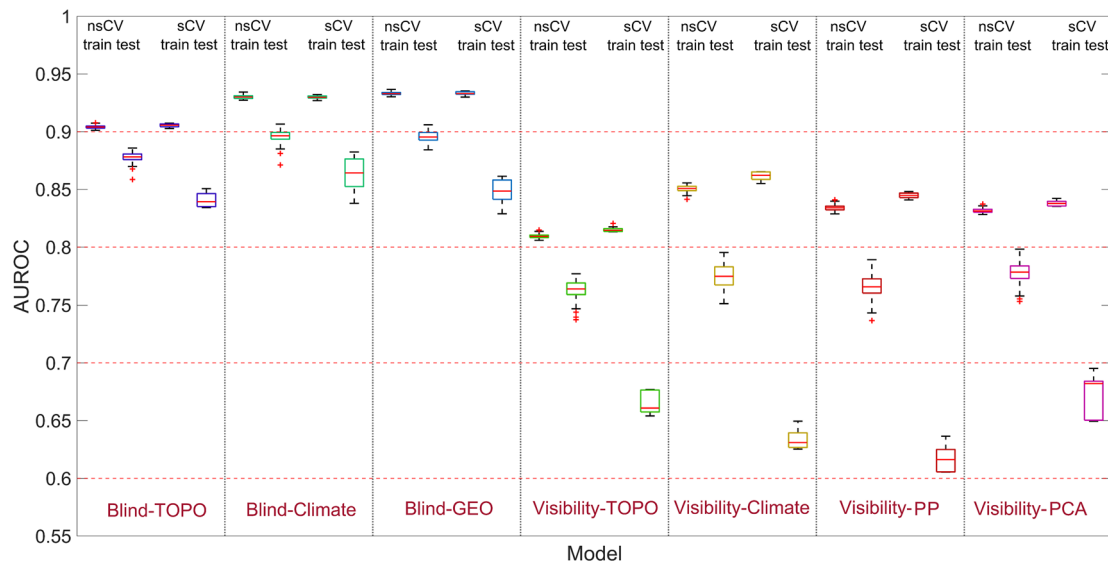
Model ID	Predictor	Penalization frequency (edf < 0.7) [%]		mDD%	Concurvity
		nsCV	sCV		
Visibility-PCA	Elevation	0.0	0.0	31.1	0.75
	Slope	0.0	0.0	18.8	0.19
	Profile curvature	0.0	0.0	16.3	0.13
	Northness	90.8	79.8	13.9	0.30
	PC1	0.2	6.0	11.8	0.73
	PC2	4.2	17.2	8.1	0.40

Map VAR-3 (Fig. 13c) reports the difference between the models visibility-PP and blind-CLIMATE; thus, it stresses the role of the inventory bias in defining the susceptibility spatial pattern. The highest negative susceptibility variations ( $< -0.6$ ), individuating areas where the susceptibility was higher for the blind-CLIMATE model than for the visibility-PP model, are located almost exclusively along the valley bottoms. As the bias-describing behaviour of altitude is smoothed through the visibility approach, negative variations characterize the valley bottoms, whereas positive variations occupy slopes located at high altitudes and in remote secondary valleys.

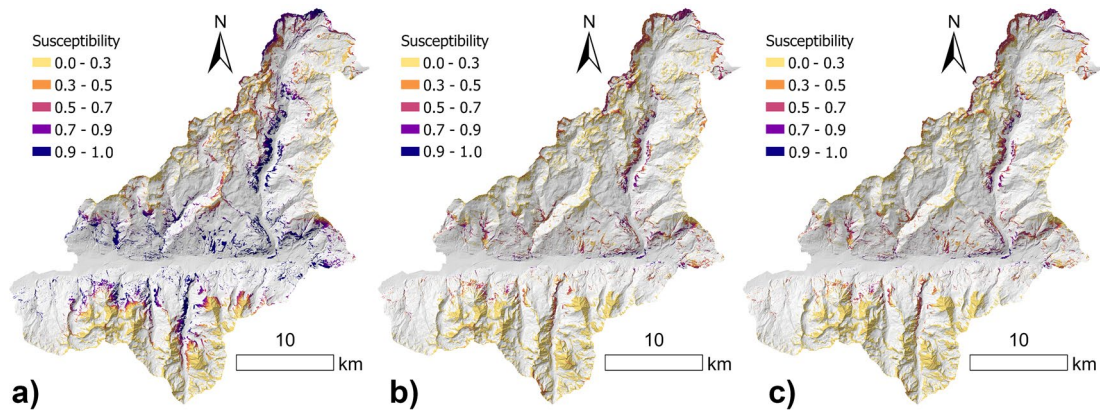
### Discussion and future perspectives

The work was aimed at translating the *non-ordinary* climatic processes and the correspondent empirical thresholds developed in the work of Bajni et al. (2021) in synthetic climatic variables to be used as predictors in a statistically based rockfall susceptibility model.

Although recently several authors (e.g. Gassner et al. 2015; Nhu et al. 2020; Bordoni et al. 2021) suggested various methods to include climatic predictors in susceptibility models, most of them were considered not optimal for the present study. The suggested methods include either the use of variables expressing average behaviours (e.g. mean annual and monthly rainfall, Broeckx et al. 2018), or responses to specific events (e.g. cumulated rainfall in specific periods of time, Knevels et al. 2020). More complex approaches dealt with the coupling of spatial susceptibility models with contingency matrices expressing rainfall thresholds exceedances (Segoni et al. 2015, 2018) or with temporal probability maps including antecedent cumulated rainfall and soil saturation degree (Bordoni et al. 2021) or by introducing synthetic variables expressing the return period of different rainstorms (e.g. Catani et al. 2013). The major limitation of the applicability of these methods to the present study consists in the need of a prior knowledge of critical amounts and durations or operational thresholds defined by Civil Protection offices, missing for this study. It was then chosen to apply an approach based on the exceedance frequency of ID type thresholds as in Dikshit et al. (2020) and Camera et al. (2021). Following the definition of *non-ordinary* conditions given in Bajni et al. (2021), the work 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. It consisted in calculating



**Fig. 11** Boxplots showing the AUROC estimate (both on train and test sets) variability observed during the non-spatial and spatial cross-validation phases for all the tested models



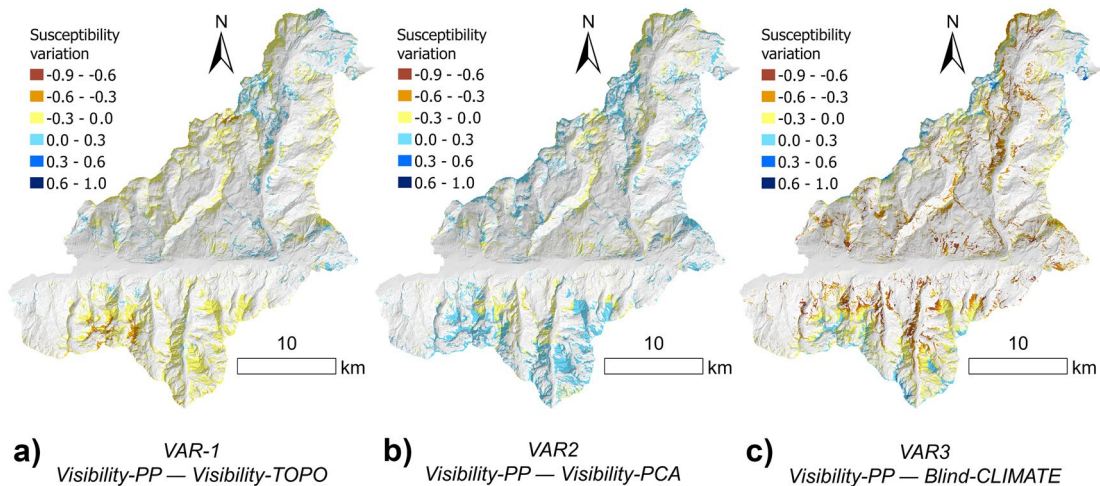
**Fig. 12** Reclassified rockfall susceptibility maps obtained from models: **a** blind-CLIMATE, **b** visibility-PP and **c** visibility-PCA

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. Different thresholds were tested for each index in Bajni et al. (2021); in particular, the threshold selected for the calculation of the mean annual exceedance frequency was the one normalized to RDN (rainy day normal, Bajni et al. 2021), as this normalization parameter is considered more suitable than MAP (mean annual precipitation) 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).

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.). This permafrost activity could be reasonably assimilated to a highly effective action of freeze–thaw cycles. However, they did not include freeze–thaw

cycles directly as a variable in the model. In addition, even if wet and dry cycles are undoubtedly recognized as a weakening process in rock masses, they are usually mainly investigated at the laboratory scale (e.g. Van der Hoven et al. 2003; Torres-Suarez et al. 2014; Zhou et al. 2017; Yang et al. 2018, 2019).

The majority of the landslide susceptibility studies available in literature assess the susceptibility model quality based only on quantitative performance. Any problem concerning the positional accuracy and spatial representativeness of the inventory is often disregarded (Reichenbach et al. 2018; Steger et al. 2017; 2021). 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, a bias towards an overrepresentation of damage-related events and an underrepresentation of remote areas is almost unavoidable. This is a long-lasting known problem in inventories obtained from public administrations (Guzzetti et al. 1999), which is nonetheless inherent to the objectives behind their compilation (Civil Protection purposes), not necessarily coinciding



**Fig. 13** Difference maps for **a** VAR-1, visibility-PP—visibility-TOPO; **b** VAR2, visibility-PP—visibility-PCA; **c** VAR-3, visibility-PP—blind-CLIMATE

with 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. The “Blind” models developed as the first step of the modelling procedure suggested an excellent rockfall discrimination capability in terms of performance ( $0.8 \leq \text{AUROC} \leq 0.9$ ). However, if performance scores are analysed together with predictors behaviour and susceptibility spatial patterns, they confirmed the existence of a bias in the inventory. Associated models showed implausible relationships (CSF curves) between rockfall susceptibility and predictors’ values (e.g. elevation, EW, WD and SWEep), if compared with the current knowledge on the physical processes leading to instability, and unlikely susceptibility spatial patterns. Similar results were obtained in other works where inventory bias issues were noticed (e.g. Steger et al. 2017, 2021; Bajni et al. 2022). 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 they might act as bias-reinforcements.

The “visibility” mask approach proposed by Knevels et al. (2020) and Bornaetxea et al. (2018) and adopted in this study proved to be valuable for reducing model inconsistencies related to the inventory bias. It needs to be noted that in the visibility-related configurations, the susceptibility values predicted outside the visibility mask correspond to areas not covered by sampling (training) points. Thus, the results over these areas may suffer from extrapolation, especially for those areas presenting predictors values beyond those of the training points. A complementary approach, which has the advantage to avoid extrapolation, is the recent work of Cignetti et al. (2021). In this study, the authors modelled rockfall susceptibility along the regional road network of Aosta Valley, inside a buffer of 250 m from roads. An additional step forward was taken by Alvioli et al. (2021), who recognized rockfall potential source areas and modelled expected consequent trajectories along the Italian railway network. Susceptibility resulted from a reclassification of trajectory counts in each cell of the modelling domain. These two studies are examples of a risk-oriented susceptibility, which has the aim to point out human infrastructures and settlements potentially subject to rockfall-related damage. However, this strategy does not match 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 represent a support for the present-time risk managing activities, but it is less suitable for a long-term planning considering the need of new sustainable development strategies in the context of climate change.

The present study confirmed that neglecting the modelled predictors’ behaviour physical plausibility, which is particularly crucial when dealing with process-driven predictors (Camera et al. 2021; Bajni et al. 2022), may lead to misleading susceptibility maps and process interpretation, even if including them would improve the quantitative performance of the model. In the optimized model (visibility-PP), EW resulted as the most important, non-penalized, physically plausible climate-related predictor, consistently with the majority of works dealing with models including cumulated precipitation-related variables (e.g. Catani et al. 2013; Capecchi et al. 2015; Kim et al. 2015; Knevels et al. 2020; Camera et al. 2021). Also, the sigmoidal shape, at low values, of its CSF was already observed in other works for variables accounting for cumulated rainfall (e.g. Tanyaş

et al. 2022). In presence of concurvity, the penalization of certain predictors 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 (Lacey et al. 2016). An example is the relationship between elevation and FT. At high altitudes ( $> 2500$  m a.s.l.), the elevation CSF shows a raising tail, which causes an increase in susceptibility over the high mountain peaks (especially in the northern part of the study area), whereas the FT predictor is penalized. This increasing susceptibility is coherent with the findings of Messenzehl et al. (2017), who found an intense rockfall activity at altitudes above 2900 m a.s.l. and imputing it to permafrost degradation causing, similarly to FT cycles, water seepage and hydrostatic pressure variations.

Finally, it is a common practice to carry out a multicollinearity (concurvity in the case of GAMs) diagnosis of predictors before the development of predictive models using both logistic regression and complex machine learning algorithms (Camera et al. 2017; Chen et al. 2017; Nohani et al. 2019; Roy et al. 2019; Yi et al. 2020; Arabameri et al. 2020; Alqadhi et al. 2021). The presence of concurvity is, however, physically reasonable, as climate processes are interconnected. As an example, EW 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 occurrence and duration of storms. Nonetheless, the application of the shrinkage option for variable selection should allow to penalize the portion of the smoothing function involved in multi-concurvity issues (Figueiras et al. 2005; Lacey et al. 2016; Bagalwa et al. 2021). This is possibly the main reason for the high penalization frequency of FT predictors in most of the models. As an alternative strategy, this study proposed to reduce climate variables to principal components (PC) and use them as predictors. Messenzehl et al. (2017; 2018) applied a similar strategy, to model rockfall activity in the Swiss Alps. The results are encouraging since the PC allowed both a reduction of concurvity and an improvement of model parsimony. In addition, the PC maintained a physically explainable behaviour. However, it needs to be noted that climatic predictors are inherently non-stationary; thus, concurvity between them (and with the other static predictors) could vary in the future, as a response to threshold exceedance pattern modifications. Also, although PC roles in the model remained physically explainable, it is difficult to discern the influence of each single climate process in defining the component behaviour and importance.

The sCV results pointed out a low spatial transferability of the models including climatic predictors, and a consistent drop of about 0.10 of the mean AUROC estimate if compared with the nsCV counterpart. Although a spatial CV is usually considered the best choice when dealing with spatial data (Brenning et al. 2012b), it may have some issues when dealing with climatic data, especially in the 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. For these reasons, a preliminary evaluation of the climatic homogeneity of the area is crucial, and an independent holdout test set, in a neighbouring region with a similar altitudinal and geographical extent, could be a better option to test the model spatial transferability.

Future studies should focus on the calculation of the rockfall-related climate indices for future periods, under different forcing conditions (i.e. share socio-economic pathways; Meinshausen et al. 2020). Rockfall susceptibility projections, possibly useful for land management and planning in the mid- to long-term period, could therefore be derived. A slight modification of the model structure, with an event-based interpretation of the predictors, could result in an instrument useful for early-warning purposes (given the availability of reliable, high-resolution weather forecasts for the area). In both cases, using cloud-based solutions (e.g. Google Earth Engine) to refine the model and process data could make possible to simulate rockfall susceptibility over any area and any scenario, assuming similar geomorphological settings (Alpine terrains). Examples of tools developed for susceptibility modelling are just starting to appear in literature (Titti et al. 2022; Wu et al. 2022).

### Conclusions

This study focused on developing and testing spatially distributed climatic predictors for rockfall susceptibility modelling. In detail, this study main outputs include:

- An analysis of the effects of not considering the rockfall data collection strategy on the model outputs;
- A strategy to deal with the inventory bias (towards roads, infrastructures and urban areas);
- A process-oriented critical analysis of GAM-associated component smoothing functions to assess the physical significance of climatic predictors;
- An assessment of the importance and roles of the climatic predictors in modelling rockfall susceptibility in the area;
- A procedure to deal with the effects of concurvity on the model outputs (by means of a principal component analysis).
- The susceptibility maps associated with the visibility-PP and visibility-PCA models resulted as valuable tools for an informed land management and local action or monitoring planning.

The selection of non-rockfall points inside a defined “visibility mask” was a valuable approach to handle the inventory bias and reduce its influence on model outputs. Also, it was fundamental to reveal the actual behaviour of predictors and assess their physical plausibility. The most important physically plausible climate-related predictors were the effective water input (EWI), the maximum snowmelt calculated over a 16-day period from snow water equivalent data (SWE<sub>max</sub>) and the wet-dry cycles (WD), which together accounted for a total mDD% of about 20%. The predictor related to the number of snowmelt events in a hydrological year (SWE<sub>ep</sub>) was always represented by a physically distorted CSF, while the predictor summarizing freeze–thaw cycles (FT) was almost always penalized as its behaviour was masked by elevation. When the climate and snow-related predictors were inserted in the susceptibility model as principal components, concurvity was efficiently reduced. Nonetheless, despite the PCA-related model embody the quality of parsimony, it may suffer from a less immediate interpretability and updatability with climate change projections.

A process-related, potentially non-stationary configuration of 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 were crucial inputs for the analysis. 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 useful, not only for a scientific interest but also to provide benefits to the mountain communities through an informed future environmental planning.

### Acknowledgements

The authors would like to thank Luigi Lombardo and another anonymous reviewer for their constructive comments, which helped improving the manuscript.

### Author contribution

Conceptualization: GB, CC and TA. Methodology: GB and CC. Formal analysis and investigation: GB and CC. Writing — original draft preparation: GB. Writing — review and editing: CC and TA.

### Funding

Open access funding provided by Università degli Studi di Milano within the CRUI-CARE Agreement. Greta Bajni received support through the PhD scholarship funded by MIUR – Ministero dell’Istruzione, dell’Università e della Ricerca. The present work was co-financed through the collaboration agreement between Università degli Studi di Milano and Centro Funzionale Regione Valle d’Aosta in the framework of the Interreg France-Italy “ALCOTRA” 2014/20 (FESR) Programme.

### Availability of data and material

Landslide data used in this work are freely accessible at <http://catastodissesti.partout.it/>.

### Declarations

**Conflict of interest** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

### References

- Abdi H, Williams LJ (2010) Principal component analysis. *WIREs. Comput Statistics* 2:433–459. <https://doi.org/10.1002/wics.101>
- Alqadhi S, Mallick J, Talukdar S, Bindajam AA, Van Hong N, Saha TK (2021) Selecting optimal conditioning parameters for landslide susceptibility: an experimental research on Aqabat Al-Sulbat,

- Saudi Arabia. *Environ Sci Pollut Res*. <https://doi.org/10.1007/s11356-021-15886-z>
- 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. *Nat Hazard* 10:831–841. <https://doi.org/10.5194/nhess-10-831-2010>
- Arabameri A, Chen W, Loche M, Zhao X, Li Y, Lombardo L, Cerda A, Pradhan B, Bui DT (2020) Comparison of machine learning models for gully erosion susceptibility mapping. *Geosci Front* 11:1609–1620. <https://doi.org/10.1016/j.gsf.2019.11.009>
- Argand E (1911) Les nappes des recouvrement des Alpes Pennines et leur prolongements structuraux. *Mat Carte Géol Suisse* 31:1–26
- Alvioli M, Santangelo M, Fiorucci F, Cardinali M, Marchesini I, Reichenbach P, Rossi M, Guzzetti F, Peruccacci S (2021) Rockfall susceptibility and network-ranked susceptibility along the Italian railway. *Eng Geol* 293:106301. <https://doi.org/10.1016/j.enggeo.2021.106301>
- Bagalwa RM, Chartin C, Baumgartner S, Mercier S, Syauswa M, Samba VC, Zabona MT, Karume K, Cizungu NL, Barthel M, Doetterl S, Six J, Boeckx P, Oost KV (2021) Spatial and seasonal patterns of rainfall erosivity in the Lake Kivu region: insights from a meteorological observatory network. *Prog Phys Geogr: Earth Environ* 45:866–884. <https://doi.org/10.1177/0309133211001793>
- Bajni G, Camera CAS, Apuani T (2021) 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 CAS, Brenning A, Apuani T (2022) Assessing the utility of regionalized rock-mass geomechanical properties in rockfall susceptibility modelling in an Alpine environment. *Geomorphol* 415:108401. <https://doi.org/10.1016/j.geomorph.2022.108401>
- Ballèvre M, Kienast JR, Vuichard JP (1986) La “nappe de la Dent Blanche” (Alpes occidentales): deux unités austroalpines indépendantes. *Eclogae Geol Helv* 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. *Clim Dyn* 55:61–75. <https://doi.org/10.1007/s00382-018-4339-4>
- Bigi G, Castellarin A, Coli M, Dal Piaz GV, Sartori R, Scandone P, Vai GB (1990) Structural model of Italy scale 1:500.000, sheet 1. C.N.R., Progetto Finalizzato Geodinamica, SELCA Firenze. <https://www.socgeol.it/438/structural-model-of-italy-scale-1-500-000.html>. Accessed 23 May 2023
- Bistacchi A, Dal Piaz GV, Massironi M, And ZM, Balestrieri ML (2001) The Aosta-Ranzola extensional fault system and Oligocene-present evolution of the north-western Alpine nappe stack. *Int J Earth Sci* 90:654–667
- Bordoni M, Vivaldi V, Lucchelli L, Ciabatta L, Brocca L, Galve JP, 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. *Nat Hazard Earth Syst Sci* 18:2455–2469. <https://doi.org/10.5194/nhess-18-2455-2018>
- Brabb EE (1984) Innovative approaches to landslide hazard and risk mapping. *Proc 4th Int Symp Landslides*. Canadian Geotechnical Society, Toronto, Canada, pp 307–324
- 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 (2012b) Improved spatial analysis and prediction of landslide susceptibility: practical recommendations. *Landslides* 9:279–290. <https://doi.org/10.1007/s10346-012-0279-2>
- 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, 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>
- Broeckx J, Vanmaercke M, Duchateau R, Poesen J (2018) A data-based landslide susceptibility map of Africa. *Earth Sci Rev* 185:102–121. <https://doi.org/10.1016/j.earscirev.2018.05.002>
- Camera CAS, 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. *Sci Total Environ* 147360. <https://doi.org/10.1016/j.scitotenv.2021.147360>
- Camera CAS, Zomeni Z, Noller JS, Zissimos AM, Christoforou IC, Bruggeman A (2017) 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>
- Capecchi V, Perna M, Crisci A (2015) Statistical modelling of rainfall-induced shallow landsliding using static predictors and numerical weather predictions: preliminary results. *Nat Hazards Earth Syst Sci* 15:75–95. <https://doi.org/10.5194/nhess-15-75-2015>
- Carrara A, Cardinali M, Detti R, Guzzetti F, Pasqui V, Reichenbach P (1991) GIS techniques and statistical models in evaluating landslide hazard. *Earth Surf Proc Land* 16(5):427–445
- Catani F, Lagomarsino D, Segoni S, Tofani V (2013) Landslide susceptibility estimation by random forests technique: sensitivity and scaling issues. *Nat Hazard* 13:2815–2831. <https://doi.org/10.5194/nhess-13-2815-2013>
- Chen W, Pourghasemi HR, 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>
- Cignetti M, Godone D, Bertolo D, Paganone M, Thuegaz P, Giordan D (2021) Rockfall susceptibility along the regional road network of Aosta Valley region (northwestern Italy). *J Maps* 17(3):54–64. <https://doi.org/10.1080/17445647.2020.1850534>
- Corò D, Galgaro A, Fontana A, Carton A (2015) A regional rockfall database: the Eastern Alps test site. *Environ Earth Sci* 74:1731–1742. <https://doi.org/10.1007/s12665-015-4181-5>
- Crozier MJ (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>
- 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. *Int J Earth Sci* 90:668–684. <https://doi.org/10.1007/s005310000177>
- Dikshit A, Sarkar R, Pradhan B, Jena R, Drukpa D, Alamri AM (2020) Temporal probability assessment and its use in landslide susceptibility mapping for Eastern Bhutan. *Water* 12:267. <https://doi.org/10.3390/w12010267>
- Duvillard P-A, Ravanel L, Deline P (2015) Risk assessment of infrastructure destabilisation due to global warming in the high French Alps. *J Alpine Res | Rev Geogr Alp* 103-2. <https://doi.org/10.4000/rga.2896>
- Ellero A, Loprieno A (2017) Nappe stack of Piemonte-Ligurian units south of Aosta Valley: new evidence from Urtier Valley (Western Alps). *Geol J* 53:1665–1684. <https://doi.org/10.1002/gj.2984>
- Figueiras A, Roca J, Cadarso-Suárez C (2005) A bootstrap method to avoid the effect of concurrency in generalised additive models in time series studies of air pollution. *J Epidemiol Community Health* 59:881–884. <https://doi.org/10.1136/jech.2004.026740>
- Filippa G, Cremonese E, Galvagno M, Isabellon M, Bayle A, Choler P, Carlson BZ, Gabellani S, Morra di Cella U, Migliavacca M (2019) Climatic drivers of greening trends in the Alps. *Remote Sens* 11. <https://doi.org/10.3390/rs11212527>

- 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>
- Frey M, Desmons J, Neubauer F (1999) The new metamorphic map of the Alps. *Schweiz Mineral Petrogr Mitt* 79:1–230
- Gariano SL, Guzzetti F (2016) Landslides in a changing climate. *Earth Sci Rev* 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; 1:429–433
- Giusti F, Dal Piaz G, Massironi M, Schiavo A (2004) Carta geotettonica della Valle d'Aosta. *Memorie Di Scienze Geologiche* 55(2003):129–149
- Gobiet A, Kotlarski S, Beniston M, Heinrich G, Rajczak J, Stoffel M (2014) 21st century climate change in the European Alps—a review. *Sci Total Environ* 493:1138–1151. <https://doi.org/10.1016/j.scitotenv.2013.07.050>
- Goetz JN, Brenning A, Petschko H, Leopold P (2015) Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. *Comput Geosci* 81:1–11
- Grämiger LM, Moore JR, Gischig VS, Loew S (2018) Thermomechanical stresses drive damage of Alpine valley rock walls during repeat glacial cycles. *J Geophys Res Earth Surface* 123:2620–2646. <https://doi.org/10.1029/2018JF004626>
- 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](https://doi.org/10.1016/S0169-555X(99)00078-1)
- 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). *Int J Rock Mech Min Sci* 42:331–349. <https://doi.org/10.1016/j.ijrmmms.2004.11.003>
- 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>
- Hoek E (2000) Analysis of rockfall hazards. *Practical Rock Eng* 117–136
- 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>
- Hosmer DW, Lemeshow S, Sturdivant RX (2013) *Applied logistic regression*. 1st ed, Wiley Series in Probability and Statistics. Wiley. <https://doi.org/10.1002/9781118548387>
- 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>
- Kim HG, Lee DK, Park C, Kil S, Son Y, Park JH (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>
- 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>
- Lacey JP, 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. *Hydrol Earth Syst Sci* 20:2467–2482. <https://doi.org/10.5194/hess-20-2467-2016>
- Li L, Cheng S, Wen Z (2021) Landslide prediction based on improved principal component analysis and mixed kernel function least squares support vector regression model. *J Mt Sci* 18:2130–2142. <https://doi.org/10.1007/s11629-020-6396-5>
- Loche M, Scaringi G, Yunus AP, Catani F, Tanyas H, Frodella W, Fan X, Lombardo L (2022) Surface temperature controls the pattern of post-earthquake landslide activity. *Sci Rep* 12:9888. <https://doi.org/10.1038/s41598-022-04992-8>
- Lombardo L, Opitz T, Ardizzone F, Guzzetti F, Huser R (2020) Space-time landslide predictive modelling. *Earth-Sci Rev* 209:103318. <https://doi.org/10.1016/j.earscirev.2020.103318>
- Magnin F, Westermann S, Pogliotti P, Ravelin L, Deline P, Malet E (2017) Snow control on active layer thickness in steep alpine rock walls (Aiguille du Midi, 3842m.a.s.l., Mont Blanc massif). *CATENA* 149:648–662. <https://doi.org/10.1016/j.catena.2016.06.006>
- Mancini M, Weindorf DC, Silva SHG, Chakraborty S, dos Teixeira AFS, Guilherme LRG, Curi N (2019) Parent material distribution mapping from tropical soils data via machine learning and portable X-ray fluorescence (pXRF) spectrometry in Brazil. *Geoderma* 354:113885. <https://doi.org/10.1016/j.geoderma.2019.113885>
- Meinshausen M, Nicholls ZRJ, Lewis J, Gidden MJ, Vogel E, Freund M, Beyerle U, Gessner C, Nauels A, Bauer N, Canadell JG, Daniel JS, John A, Krummel PB, Luderer G, Meinshausen N, Montzka SA, Rayner PJ, Reimann S, Smith SJ, van den Berg M, Velders GJM, Vollmer MK, Wang RHJ (2020) The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500. *Geosci Model Dev* 13:3571–3605. <https://doi.org/10.5194/gmd-13-3571-2020>
- 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 Periglac Process* 29:135–151. <https://doi.org/10.1002/ppp.1976>
- 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. *Rock Mech Rock Eng* 55:2149–2172. <https://doi.org/10.1007/s00603-022-02786-y>
- Nhu V-H, Shirzadi A, Shahabi H, Chen W, Clague JJ, Geertsema M, Jaafari A, Avand M, Miraki S, Talebpoor Asl D, Pham BT, Ahmad BB, 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>
- 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. *Theoret Appl Climatol* 131:1479–1491. <https://doi.org/10.1007/s00704-017-2066-0>
- Nohani E, Moharrami M, Sharafi S, Khosravi K, Pradhan B, Pham BT, Lee S, Melesse MA (2019) Landslide susceptibility mapping using different GIS-based bivariate models. *Water* 11. <https://doi.org/10.3390/w11071402>
- 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. *Nat Hazard* 12:2283–2298. <https://doi.org/10.5194/nhess-12-2283-2012>
- Orusa T, Borgogno Mondino E (2021) Exploring short-term climate change effects on rangelands and broad-leaved forests by free satellite data in Aosta Valley (Northwest Italy). *Climate* 9:47. <https://doi.org/10.3390/cli9030047>
- 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. *Theoret Appl Climatol* 137:1765–1784. <https://doi.org/10.1007/s00704-018-2673-4>
- Pereira S, Zêzere JL, Bateira C (2012) Technical note: assessing predictive capacity and conditional independence of landslide predisposing factors for shallow landslide susceptibility models. *Nat Hazard* 12:979–988. <https://doi.org/10.5194/nhess-12-979-2012>

- Pignone F, Reborà N (2014) GRISO: rainfall generator of spatial interpolation events in regional climate simulations for Europe and the Alpine region. *J Geophys Res: Atmos* 118:3610–3626. <https://doi.org/10.1002/jgrd.50297>
- Rajczak J, Pall P, Schär C (2013) Projections of extreme precipitation events in regional climate simulations for Europe and the Alpine region. *J Geophys Res: Atmos* 118:3610–3626. <https://doi.org/10.1002/jgrd.50297>
- Ravanel L, Deline P (2010) Climate influence on rockfalls in high-Alpine steep rockwalls: the north side of the Aiguilles de Chamoni (Mont Blanc massif) since the end of the ‘Little Ice Age.’ *Holocene* 21:357–365. <https://doi.org/10.1177/0959683610374887>
- Reichenbach P, Rossi M, Malamud BD, Mihir M, Guzzetti F (2018) A review of statistically-based landslide susceptibility models. *Earth Sci Rev* 180:60–91. <https://doi.org/10.1016/j.earscirev.2018.03.001>
- Roberts DR, Bahn V, Ciuti S, Boyce MS, Elith J, Guillerà-Arroita G, Hauenstein S, Lahoz-Monfort JJ, Schröder B, Thuiller W, Warton DI, Wintle BA, Hartig F, Dormann CF (2017) Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* 40:913–929. <https://doi.org/10.1111/ecog.02881>
- Roy J, Saha S, Arabameri A, Blaschke T, Bui DT (2019) A novel ensemble approach for landslide susceptibility mapping (LSM) in Darjeeling and Kalimpong Districts, West Bengal, India. *Remote Sens* 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
- Sabokbar HF, Roodposhti MS, Tazik E (2014) Landslide susceptibility mapping using geographically-weighted principal component analysis. *Geomorphology* 226:15–24. <https://doi.org/10.1016/j.geomorph.2014.07.026>
- 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>
- 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, Tofani V, Rosi A, Catani F, Casagli N (2018) Combination of rainfall thresholds and susceptibility maps for dynamic landslide hazard assessment at regional scale. *Front Earth Sci* 6:85. <https://doi.org/10.3389/feart.2018.00085>
- Steger S, Brenning A, Bell R, Petschko H, Glade T (2016) 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, 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, 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. *Sci Total Environ* 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, Tiranti D, Huggel C (2014) Climate change impacts on mass movements — case studies from the European Alps. *Sci Total Environ* 493:1255–1266. <https://doi.org/10.1016/j.scitotenv.2014.02.102>
- Tanyaş H, Hill K, Mahoney L, Fadel I, Lombardo L (2022) 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. *Eng Geol* 297:106504. <https://doi.org/10.1016/j.enggeo.2021.106504>
- Tartarotti P, Guerini S, Rotondo F, Festa A, Balestro G, Bebout GE, Cannò E, Epstein GS, 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>
- Titti G, Napoli GN, Conoscenti C, Lombardo L (2022) Cloud-based interactive susceptibility modeling of gully erosion in Google Earth Engine. *Int J Appl Earth Obs Geoinf* 115:103089. <https://doi.org/10.1016/j.jag.2022.103089>
- Torres-Suarez MC, 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. *J Rock Mech Geotech Eng* 6:257–268. <https://doi.org/10.1016/j.jrmge.2014.04.004>
- Triglia A, Frattini P, Casagli N, Catani F, Crosta G, Esposito C, Iadanza C, Lagomarsino D, Mugnozza GS, Segoni S (2013) Landslide susceptibility mapping at national scale: the Italian case study. *Landslide Science and Practice*. Springer, pp 287–295
- Van der Hoven SJ, Solomon DK, Moline GR (2003) Modeling unsaturated flow and transport in the saprolite of fractured sedimentary rocks: effects of periodic wetting and drying. *Water Resour Res* 39. <https://doi.org/10.1029/2002WR001926>
- van Westen CJ, Castellanos E, Kuriakose SL (2008) Spatial data for landslide susceptibility, hazard, and vulnerability assessment: an overview. *Eng Geol* 102:112–131. <https://doi.org/10.1016/j.enggeo.2008.03.010>
- Wood SN (2017) Generalized additive models: an introduction with R, 2nd edn. Chapman and Hall/CRC, Boca Raton, FL, USA
- World Meteorological Organization (1989) Calculation of monthly and annual 30-year standard normals. WCDP-No. 10, WMOTD/No. 341, World Meteorological Organization, Washington D.C., USA, p 11. [https://library.wmo.int/index.php?lvl=notice\\_display&id=11642](https://library.wmo.int/index.php?lvl=notice_display&id=11642). Accessed 23 May 2023
- World Meteorological Organization (2007) The role of climatological normals in a changing climate. WCDMP-No. 61, WMOTD/No. 1377, World Meteorological Organization, Geneva, Switzerland, p 130. [https://library.wmo.int/index.php?lvl=notice\\_display&id=16659](https://library.wmo.int/index.php?lvl=notice_display&id=16659). Accessed 23 May 2023
- Wu W, Zhang Q, Singh VP, Wang G, Zhao J, Shen Z, Sun S (2022) A data-driven model on Google Earth Engine for landslide susceptibility assessment in the Hengduan Mountains, the Qinghai-Tibetan Plateau. *Remote Sens* 14:4662. <https://doi.org/10.3390/rs14184662>
- 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. *Environ Earth Sci* 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>

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. *Eng Geol* 220:1–12. <https://doi.org/10.1016/j.enggeo.2017.01.017>

Zhu L, Wang G, Huang F, Li Y, Chen W, Hong H (2022) Landslide susceptibility prediction using sparse feature extraction and machine learning models based on GIS and remote sensing. *IEEE Geosci Remote Sens Lett* 19:1–5. <https://doi.org/10.1109/LGRS.2021.3054029>

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10346-023-02091-x>.

---

**Greta Bajni · Corrado A. S. Camera** (✉) · **Tiziana Apuani**

Dipartimento di Scienze della Terra “A. Desio”, Università degli Studi di Milano, Via Mangiagalli 34, Milan 20133, Italy

Email: [corrado.camera@unimi.it](mailto:corrado.camera@unimi.it)