INTRODUCING INTENSE RAINFALL AND SNOWMELT VARIABLES TO IMPLEMENT A PROCESS-RELATED NON-STATIONARY SHALLOW LANDSLIDE SUSCEPTIBILITY ANALYSIS

Corrado A.S. Camera[1*], Greta Bajni[1], Irene Corno[1], Mattia Raffa[1], Stefania Stevenazzi[1], Tiziana Apuani[1]

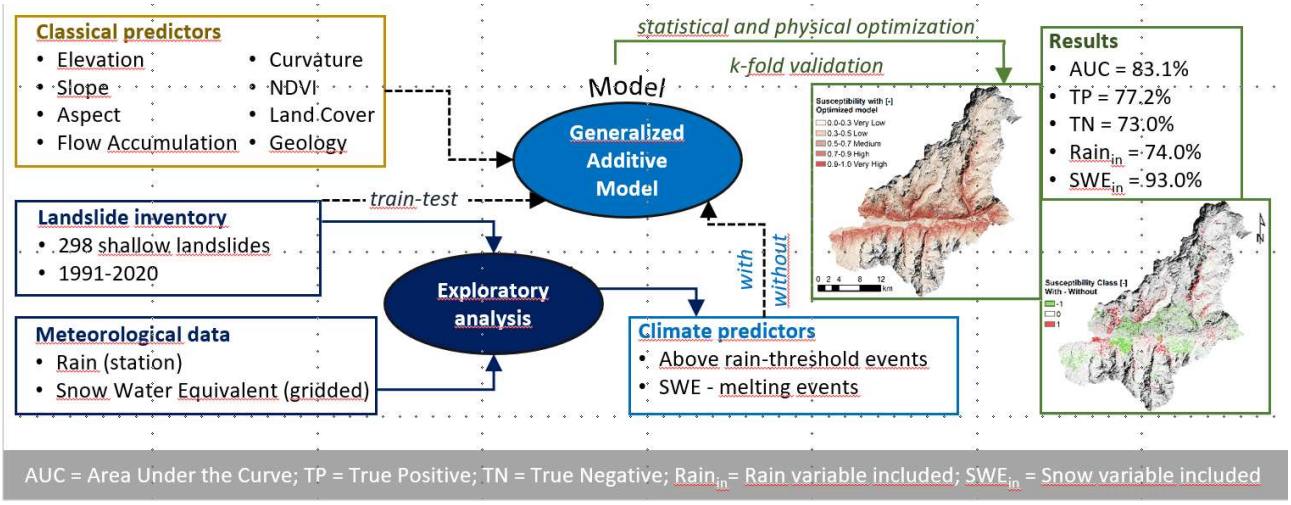
[1] *Dipartimento di Scienze della Terra "A. Desio", Università degli Studi di Milano, Milan (Italy)*
[*] *Corresponding author email and telephone: corrado.camera@unimi.it, +39 0250315548*

Highlights

- Intense rainfall and snowmelt were included in a landslide susceptibility analysis
- The statistical and physical importance of the variables were verified by a GAM
- The climate related variables together explained 5% of the model deviance
- The climate variables modified the susceptibility class in 11% of the study area
- These variables allow to explore climate change impacts on landslide susceptibility

Keywords: Climate variables, Slides in soil, Flowslides, Generalized Additive Models, Snow Water Equivalent, Aosta Valley

Graphical Abstract



Abstract

The study objective was to derive a susceptibility model for shallow landslides that could include process-related non-stationary variables, to be adaptable to climate changes. We selected the territory of the Mont-Emilius and Mont-Cervin Mountain Communities (northern Italy) as the study area. To define summary variables related to landslide predisposing and triggering processes, we investigated the relationships between landslide occurrences and intense rainfall and snowmelt events (period 1991–2020). For landslide susceptibility mapping, we set up a Generalized Additive Model. We defined a reference model through variable penalization (relief, NDVI, land cover and geology predictors). Similarly, we optimized a model including the climate variables, checking their smooth functions to ensure physical plausibility. Finally, we validated the optimized model through a *k*-fold cross-validation and performed an evaluation based on contingency tables, area under the receiver operating characteristic curve (AUROC) and variable importance (decrease in explained variance). The climate variables that resulted as being statistically and physically significant are the effective annual number of rainfall events with intensity–duration characteristics above a defined threshold (EAT_{ean}) and the average number of melting events occurring in a hydrological year (ME_n). In the optimized model, EAT_{ean} and ME_n accounted for 5% of the explained deviance. Compared to the reference model, their introduction led to an increase in true positive rate and AUROC of 2.4% and 0.8%, respectively. Also, their inclusion caused a transition of the vulnerability class in 11.0% of the study area. The *k*-fold validation confirmed the statistical significance and physical plausibility of the meteorological variables in 74% (EAT_{ean}) and 93% (ME_n) of the fitted models. Our results demonstrate the validity of the proposed approach to introduce process-related, non-stationary, physically-plausible climate variables within a shallow landslide susceptibility analysis. Not only do the variables improve the model performance, but they make it adaptable to map the future evolution of landslide susceptibility including climate changes.

1 Introduction

Landslides are amongst the natural hazards globally causing the highest economic damage and claiming the largest number of victims every year (Herath and Wang, 2009; Kjekstad and Highland, 2009; Petley, 2012; UNISDR, 2015). The capacity to predict these events, or at least to recognize those areas most prone to failure, is therefore beneficial for land use management, infrastructure planning, increased resilience in communities, and disaster risk reduction (e.g., Camera et al., 2015; Dou et al., 2019; Mertens et al., 2016).

According to Brabb (1984), landslide susceptibility is defined as the likelihood of a mass movement occurring in an area based on its characteristics. Landslide susceptibility is therefore the spatial component of the hazard, which in addition includes a quantification of the frequency and magnitude of an event (Guzzetti et al., 2005). Landslide susceptibility is therefore valuable for land and environmental management but simpler to assess than hazard. Also, it is very suitable for investigation through quantitative statistical models, which can identify relationships between the distribution of previously occurred landslides and environmental factors (Guzzetti et al., 2006). As in other fields (e.g., groundwater vulnerability and digital soil mapping; Camera et al., 2017; Stevenazzi et al., 2015), in the last two decades the application of such approaches for mapping purposes has grown continuously (Carrara and Pike, 2008; Reichenbach et al., 2018). The reasons are the increasing number of data-mining algorithms and GIS tools available, and the release of several regional and global remote-sensing datasets (Minasny and McBratney, 2016). Recently, Reichenbach et al. (2018) carried out an extensive review on the use of statistically-based models for landslide susceptibility analysis, which the authors concluded by furnishing guidelines to perform a high-quality assessment and highlighting open research questions in the field. Amongst the latter, the unknown impacts of climate change on landslide susceptibility were pointed out.

Indeed, in mountainous environments the occurrence of landslides has strict links with climate-related processes such as intense rainfall and snowmelt (Luino et al., 2020; Schilirò et al., 2021), which are both expected to vary in the future due to global warming (Beniston et al., 2018; Gobiet et al., 2014). In particular, intense rainfall is often recognized as an immediate cause (i.e., a triggering factor) for the initiation of shallow landslides (Di Napoli et al., 2021), while the effect of snowmelt is reported to be both immediate and cumulative (i.e., snowmelt can be considered both a triggering and a predisposing factor - Lucas et al., 2020; Stumvoll et al., 2020; Subramanian et al., 2020). In terms of their expected changes, for north-western Italy, from high resolution (2.2 km) climate simulations forced under the RCP 8.5 (IPCC, 2014), Ban et al. (2020) showed increases in the 10-year return value of 5-day, 1-day and 1-hour precipitation up to 30% (2081–2090 vs 1991–2000, highest values in winter). For snow dynamics, projected changes include variations in the rain-to-snow fraction, a seasonal shift of snowmelt and a general decrease in the length of the snow cover season (Beniston et al., 2018; Citrini et al., 2020; Javadinejad et al., 2020). These changes in landslide driving factors are reasonably expected to modify the response of mountainous environments to instability. However, their impacts still need to be completely understood (increase or decrease) and quantified in terms of landslide susceptibility (Gariano and Guzzetti, 2016).

Until 2016, only 2.8% of the studies on landslide susceptibility included a precipitation-related variable and as few as 0.3% of them added other climate predictors (Reichenbach et al., 2018). When included, precipitation is often in the form of its average characteristics, such as mean annual rainfall,

mean monthly rainfall and rainy days normal (e.g., Broeckx et al., 2018; Chen and Li, 2020; Fang et al., 2020; Nahayo et al., 2019; Nhu et al., 2020), thus failing to consider extreme processes that may have a high influence on the occurrence of landslides. In addition, over the Alps the projected future changes in mean precipitation differ significantly from changes in intense events (Ban et al., 2020). Efforts to consider intense rainfall mainly concern event-based susceptibility studies (e.g., Knevels et al., 2020) and those targeting the development of early warning systems (e.g., Bordoni et al., 2020b). For early warning, classical susceptibility models based on stationary variables (e.g., relief, geology) are coupled with information on intense rainfall external to the model, which could be in the form of a rainfall threshold for landslide occurrence (Segoni et al., 2015, 2018b) or of an additional temporal statistical model (Bordoni et al., 2020b). In event-based susceptibility models, authors (e.g., Kim et al., 2015; Knevels et al., 2020) usually introduce variables representing multiple-day maximum cumulated precipitation and select critical durations through model optimization (i.e., selecting durations maximizing the model performance indices). Some authors (Gassner et al., 2015; Kim et al., 2015; Shou and Lin, 2020) also used the optimized model to derive future scenarios of landslide susceptibility calculating the appropriate climate variables from climate simulations. However, the weather conditions that could lead to the occurrence of landslides over a certain area are various in terms of the cumulated rainfall, duration and intensity (Peruccacci et al., 2017), therefore event-based approaches and the related variables are not able to completely capture them. Regarding snowmelt, to our knowledge no study has so far tried to summarize this process into a variable for landslide susceptibility assessment. Therefore, a comprehensive set of variables is still needed for a thorough inclusion of intense rainfall and snowmelt in landslide susceptibility analysis.

Another open question in landslide susceptibility analysis is the physical and geomorphic plausibility of the optimized statistical model (Steger et al., 2016). Models are evaluated quantitatively through performance indices, but different models with different outputs could return equally good predictive performances (Sterlacchini et al., 2011). For this reason, it is preferable to use modelling algorithms that are easy to interpret, which also allows a verification of the physical plausibility of the role played by each variable included in the analysis.

The aim of this study is to develop a statistically-based model for the definition of the spatial susceptibility to shallow landslides, whose occurrence relates to excess pore water pressure (slides in soil, sand/silt/debris flowslides and mud flows, as defined by Hungr et al., 2014), including a comprehensive set of process-related non-stationary variables that are adaptable to reveal climate change impacts. As the study area, we focused on the Mountain Communities of Mont Cervin and Mont Emilius, within the Aosta Valley Region (northern Italy). Specific objectives of the study consisted in: i) recognizing relationships between intense rainfall, snowmelt and shallow landslide events; ii) summarizing the recognized relationships in variables to be used in a statistical model; iii) evaluating the statistical significance and physical plausibility of the climate variables within a generalized additive model; iv) quantifying the added value of the climate variables in terms of model fitting and predictive performance.

In the following section we introduce the main geographical and geological characteristics of the study area as well as the landslide inventory and the meteorological data available. In Section 3, we describe the exploratory analysis that we conducted for the definition of summary variables relating climate (intense rainfall and snowmelt) to landslide occurrence and how we included them in a

statistically-based landslide susceptibility model. Section 4 presents the results of the exploratory analysis and those related to the statistical model evaluation, validation and interpretation. In Section 5 we discuss the main results against previous literature, while Section 6 summarizes the most critical findings.

2 Study area and available data

2.1 Study area

The study focuses on the Mountain Communities of Mont Cervin and Mont Emilius, located in the central part of Aosta Valley (northern Italy), in the western sector of the Alpine chain (Fig. 1, Fig. 2). The study area extends for 670 km^2. The elevation ranges from about 400 m a.s.l. in the main E–W valley bottom, to the 4478 m a.s.l. of the Matterhorn (Mont Cervin), dominating Valtournanche and the village of Cervinia.

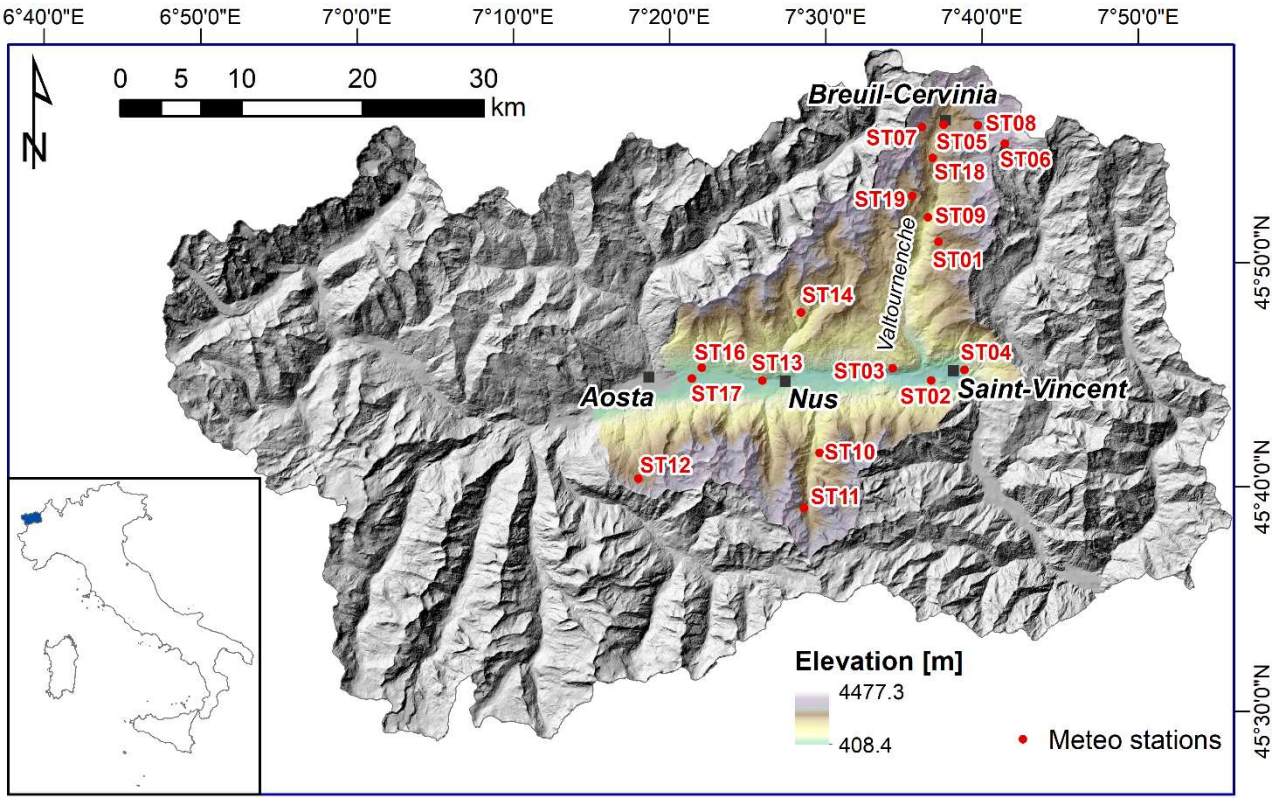


Fig. 1. Geographical setting of the study area, including orography, main towns and meteorological stations. For meteorological stations IDs see Table 1.

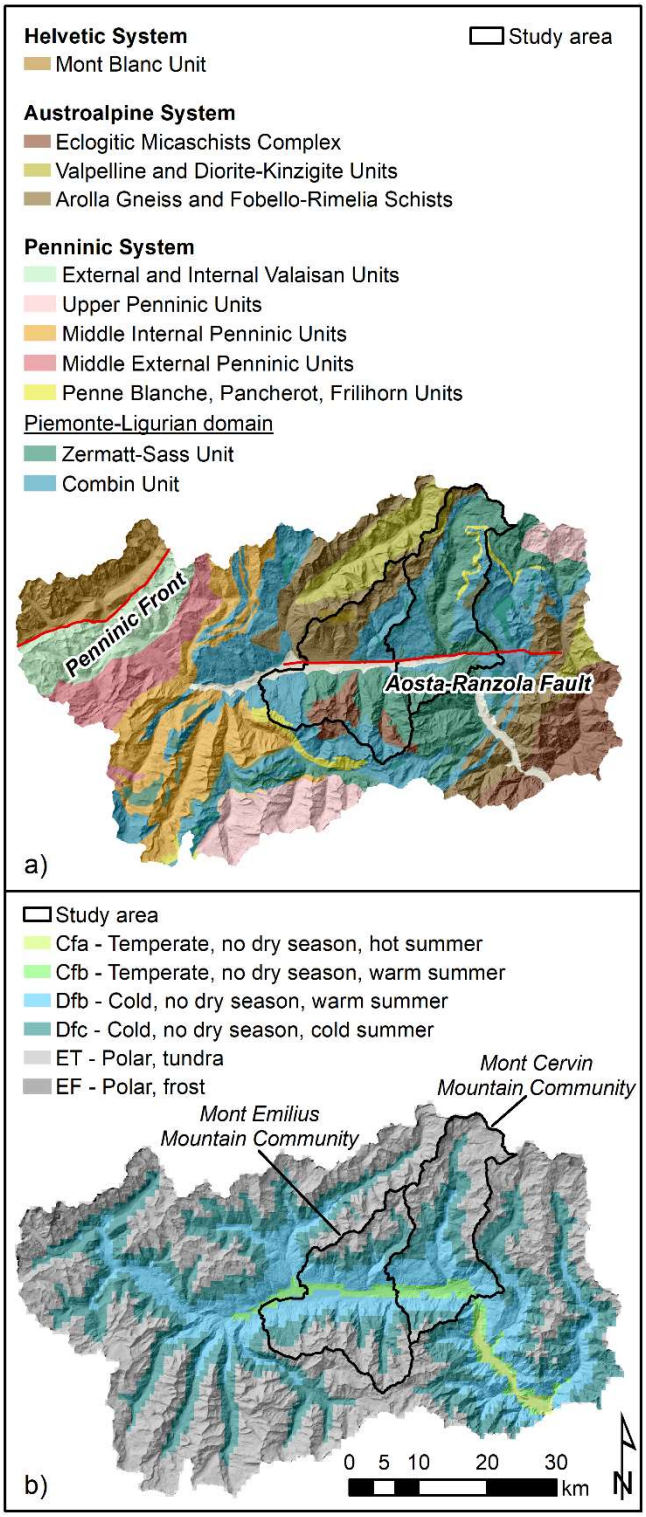


Fig. 2. a) Simplified geological map of Aosta Valley and b) dominant climate in the region (from Beck et al., 2018).

The topographical characteristics of the area are the result of a complex geological and geomorphological history. The Aosta Valley shows evidence of the orogenic processes related to the Alpine collision and the neo-tectonic brittle deformation phase that followed, leading to the development of the Aosta–Ranzola fault system (Dal Piaz et al., 2003). The latter is the tectonic element along which the main E–W axis of Aosta Valley formed. The study area lies in the Axial

Belt of the chain, i.e., in the core of the Western Alps between the Insubric Line to the east and the Penninic Front to the west (Fig. 2a). The dominant units in the area are those related to the Piemonte–Ligurian oceanic domain (Penninic System) and the Austroalpine outlier (Ellero and Loprieno, 2018). The main lithologies consist of metasediments (calcschists, marbles and quartzites) and metabasites (serpentinites, amphibolites, metagabbros and prasinites) belonging to the Piemonte–Ligurian domain and various associations of rocks of different metamorphic facies (granulitic, greeenschists, eclogitic) belonging to the Austroalpine domain (Dal Piaz et al., 2010). Besides the structural–geological processes, the glacial dynamics, particularly active during the Pleistocene, influenced the evolution of the relief too (Ratto et al., 2003). Glacial processes are indirectly conditioning the actual slope dynamics, including gravitational phenomena, principally due to the delayed action of debuttressing caused by glacier retreat (Giordan et al., 2018).

The difference in elevation is reflected in the climate characteristics. According to Koppen's climate classification for the period 1980–2016 (Beck et al., 2018), the valley bottoms present a temperate climate with warm summers, while moving towards higher elevations a transition occurs to a cold climate with warm to cold summers, till a tundra or frost-like climate on top of the highest peaks (Fig. 2b). The common trait is the absence of a dry season. Bajni et al. (under review) recognized a bimodal distribution of precipitation throughout the study area (reference period 1990–2018), with peaks in spring and autumn (generally May and November). The northern head of Valtournenche is found to be the wettest part of the study area (average annual rainfall around 960 mm and annual average snow intakes around 690 cm). In addition, the Mont Cervin Community is more humid than the Mont Emilius Community (average annual rainfall of 810 mm and 671 mm, respectively), thus showing a decreasing trend in precipitation moving from East to West.

2.2 Landslide inventory

The Aosta Valley Region established its Landslide Inventory during the 1990s. The first version was an alphanumeric paper catalogue bringing together records collected from different archives (technical regional and municipal offices, scientific organizations, libraries, parishes) and dating back to the XVIII century. After the large flood event of October 2000 (Ratto et al., 2003), the Region started to add cartographic information to the inventory. In 2001, Aosta Valley joined also the national programme for the creation of a digital, cartographic and alphanumeric landslide events database (Inventario dei Fenomeni Franosi, IFFI project, https://www.progettoiffi.isprambiente.it/en/inventory). Within the IFFI project, the Region conducted field surveys in all the areas not affected by the October 2000 flood, reviews of additional data from archives (mainly of forestry stations) and data collection through photointerpretation. The last step included data homogenization through comparison, integration and validation of all entries with geological maps and PSinSar points.

Nowadays, the inventory is completely digital and publicly available (http://catastodissesti.partout.it/). The Regional Civil Protection Department and the Forest Corps update it continuously through regular surveys or following warnings from private citizens. Each record includes the type of the event (large flood, river-bank erosion, rockfall, debris flow, earth flows, earth slides or landslide in a broad sense), the date of occurrence, the triggering location coordinates and a form describing its main characteristics (including pictures). According to Reichenbach et al. (2018), the inventory is classified as multi-temporal.

For the present study, we extracted all the triggering locations of the events classified as earth flows and earth slides (i.e., those corresponding to the slides in soil, sand/silt/debris flowslides and mud flows of Hungr et al., 2014) that occurred in the Mountain Communities of Mont Cervin and Mont Emilius between October 1990 and September 2020. This query resulted in 298 events. In addition, we extracted the same type of events that occurred in the same areas but without information on the initiation date. These events amounted to 264. Next, we checked all points by overlaying them on orthophotos and on the land use and geological maps (see Section 3.1.3 for further details on the maps) to verify their consistency in terms of the triggering areas and materials involved. In the case of points falling in the middle of landslide scars or on rock outcrops within 100 m from loose deposits, we repositioned the points on the main scarp or on the loose terrain land use, assuming a small error in recording the location.

2.3 Meteorological data

2.3.1 Station data

In the study area, 18 meteorological stations are currently operating (Fig. 1). The Agenzia Regionale per la Protezione dell'Ambiente (ARPA, http://www.arpa.vda.it/it) and the Centro Funzionale Regione Autonoma Valle d'Aosta (CF, https://cf.regione.vda.it/portale_dati.php) manage them and make the data publicly available. This study focused on rainfall and snow height data. Most data available for the past century have a daily temporal resolution for both rainfall and snow height. Automatic stations have substituted the old manual ones starting in 2002, allowing the recording of hourly and sub-hourly data too. Table 1 presents details of the parameters acquired at each station and the length of the available time series used in this study.

For the present study, we extracted all the available hourly rainfall and daily snow height data from October 1990 to September 2020. We selected this time window to match the landslide data. Also, our intention was to base our analysis on a 30-year period, typical for carrying out a climatological analysis (e.g., Menne et al., 2012). However, the longest available hourly time series dates back to 1994. Based on the climate normal concept, we assumed that the available data could be considered representative of the 30-year period (World Meteorological Organization - WMO, 2007). For those stations at which sub-hourly data were available, we aggregated them to hourly through a simple R code (https://www.r-project.org), since it was the finest common temporal resolution.

2.3.2 Snow Water Equivalent (SWE) raster data

ARPA and CF developed a raster Snow Water Equivalent (SWE) dataset for the whole Aosta Valley Region for the period 2001 to the present (Filippa et al., 2019). The dataset 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. The map shows, cell by cell, the actual volume of water stored as snow expressed in terms of equivalent water height [m]. The dataset is the product of the Snow Cover Area (SCA, derived from satellite data), interpolated snow height data (derived from station data, additional manual measures and topographic variables), and estimated snow density data (from manual measures). Due to the interpolated snow height factor, the dataset does not guarantee the conservation of mass between consecutive SWE maps. For this study, ARPA made the whole dataset available in ascii format following a request. A map visualizer is present on the ARPA website (https://www.arpa.vda.it/it/effetti-sul-territorio-dei-cambiamenti-climatici/neve/swe).

Table 1. Meteorological stations operating in the study area, with details on the parameters measured and used in this study (R is rainfall and S is snow height), together with acquisition period and temporal resolution.

ID	*Name*	*Elev [m]*	*R*	*S*	*P-hourly from*	*S-daily from*
ST01	Chamois Lac de Lou	2020	Yes	Yes	01/01/2002	19/09/2007
ST02	Pontey ponte Dora Baltea	473	Yes	No	04/10/2007	NA
ST03	Saint-Denis Raffort	840	Yes	No	01/01/2002	NA
ST04	Saint-Vincent-Terme	626	Yes	No	25/06/2007	NA
ST05	Valtournenche Breuil Cervinia	1998	Yes	Yes	09/06/2010	09/06/2010
ST06	Valtournenche Cime Bianche	3100	No	No	NA	NA
ST07	Valtournenche Grandes Murailles	2566	Yes	No	01/05/2007	NA
ST08	Valtournenche Lago Goillet	2541	Yes	Yes	05/08/2003	01/10/1990
ST09	Valtournenche Maen	1310	Yes	No	01/04/2002	NA
ST10	Fénis Clavalité	1531	Yes	No	06/08/2003	NA
ST11	Fénis Lavodilec	2250	Yes	No	26/07/2007	NA
ST12	Gressan Pila-Leissé	2280	Yes	No	17/07/2007	NA
ST13	Nus Les Iles	534	Yes	No	22/05/2007	NA
ST14	Nus Saint Barthélemy osservatorio	1675	Yes	Yes	01/09/2002	01/04/2003
ST15-16	Pollein - Quart Ollignan (merged)	650	Yes	No	01/10/2002	NA-
ST17	Saint-Christophe-Aeroporto	545	Yes	Yes	01/01/1994	10/04/2015
ST18	Valtournenche-Perreres	1750	No	Yes	NA	01/10/1990
ST19	Lago Cignana	2170	No	Yes	NA	01/10/1990

3 Methods

The methodology consists of two main steps. First, we conducted an exploratory data analysis to investigate the links between shallow landslide initiation and meteorological processes, related to rainfall and snowmelt. The aim of this first step was to derive summary climate variables for the reference period to be used in the susceptibility analysis. Second, we introduced the summary variables into the susceptibility model and we evaluated their statistical and physical significance, as well as their added value in comparison to a classical model built on relief, land use and geological predictors. We summarize the methodological approach in Fig. 3 and describe the details of the two steps in the sub-sections that follow.

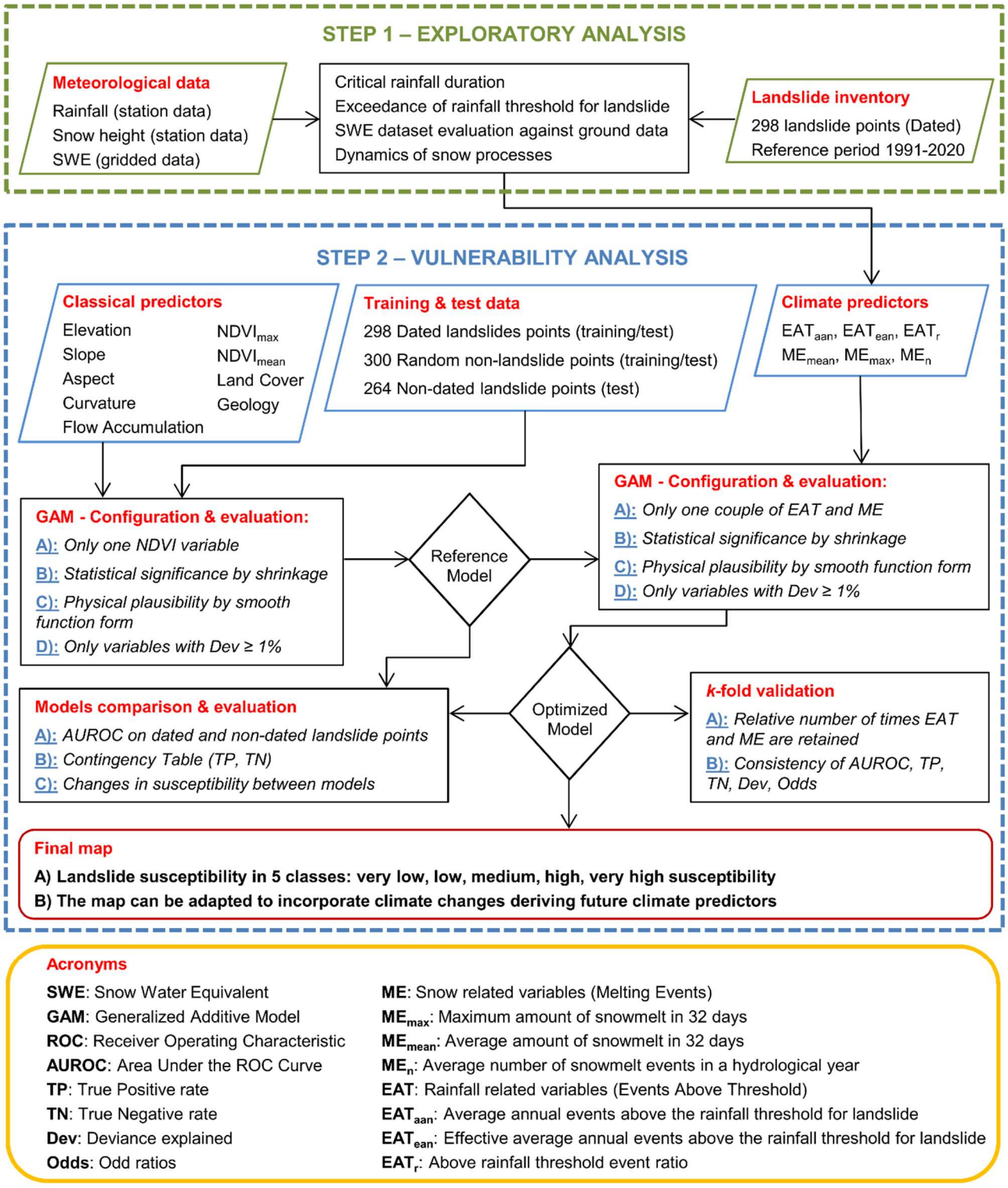


Fig. 3. Flowchart summarizing the methodological approach of the study.

3.1 Definition of model variables

3.1.1 Rainfall-related variables

We investigated the relationships between rainfall events and the triggering of shallow landslides following two approaches, which slightly modify the methodology presented in (Corno, 2019). The first approach took inspiration from the studies of Gassner et al. (2015) and Knevels et al. (2020). Developing event-based susceptibility analyses, they introduced multiple-day maximum cumulated precipitation in their statistical models and selected critical durations through model optimization (i.e., selecting durations maximizing the model performance indices). To transfer this approach to a multi-temporal analysis, we tried to identify critical durations *a priori*. We assigned the dated events to a meteorological station recording rainfall based on proximity (Voronoi polygons). We selected this approach because the number of stations, given also the very complex topography, was too limited to develop an hourly gridded dataset (Maraun and Widmann, 2018). We calculated for all stations the maximum cumulated rainfall for 1-, 3-, 6-, 12-, 24-, 48-, 72-, 96- and 120-hour periods for each month for which hourly data were available. Next, we calculated the cumulated rainfall in the various durations that occurred before the shallow landslide events and we compared it with the distribution of the monthly maxima. Since for most events we did not know the exact hour of occurrence, for each duration we used the maximum cumulated values recorded on the day of the event. Following Bajni et al. (under review), we considered a specific duration as critical if 50% of the events fell above the 75th percentile of its corresponding monthly maxima distribution. In this exploratory phase, we discarded the events attributed to a meteorological station with a time series not covering the dates on which the landslide occurred.

The second approach overlooked specific durations. We based it on rainfall thresholds for landslide occurrence (e.g., Segoni et al., 2018a). Based on the events analysed for the previous approach, we verified and adjusted the threshold proposed by Aleotti (2004) for the Piedmont region (north-western Italy). To adjust the threshold, we verified that at least three meteorological events associated with a landslide, other than the October 2000 big flood on which the threshold was built, fell over it for at least one of the examined durations. We defined an event as starting when a non-zero rainfall (i.e., a minimum of 0.2 mm) occurred and the same event ending after 24 hours without precipitation. For all recognized events, we verified if they exceeded the threshold for at least a duration between 1 and 120 hours. We attributed to each event a value of 1 or 0 for threshold exceedance or non-exceedance, respectively. From the classified (1–0) rainfall events, we calculated three indices for each station: i) the average annual number of rainfall events above the threshold (EAT_{aan}); ii) the effective annual number of rainfall events above the threshold (EAT_{ean}); iii) the above threshold event ratio (EAT_r). Spatially, we assigned the values through Voronoi polygons for consistency with the exploratory analysis. Calculation details of the three indices follow:

$$\mathrm{EAT_{aan}} = \frac{n}{years} \quad (1)$$

$$\mathrm{EAT_{ean}} = \frac{n}{(data-nas)/hours_y} \quad (2)$$

$$\mathrm{EAT_r} = \frac{n}{n_{tot}} \quad (3)$$

where n is the number of events above the threshold, n_{tot} is the total number of rainfall events recorded within a station time series, *years* is the number of hydrological years in the station time series, $hours_y$ is the number of hours in a year, *data* is the number of expected data entries in the case of a complete time series, and *nas* is the number of missing values.

3.1.2 Snow-related variables

The analysis on the snow data was two-fold. Firstly, we wanted to demonstrate that the regional SWE dataset is able to reproduce snow dynamics as recorded by ground data at the few meteorological stations. Secondly, we wanted to use snow ground data, because of its higher temporal resolution, to investigate with which dynamics and on which temporal basis snowmelt influences the initiation of shallow landslides. These two steps were necessary to verify the coherence between the datasets and to allow reliable spatial variables to be derived for the susceptibility analysis.

To overcome the issue related to the SWE dataset's lack of mass conservation at the single cell, we aggregated the SWE values in sub-basins calculating averages (additional details in Raffa, 2020). We selected a basin-based aggregation rather than a raster (squared cell) approach because we considered snowmelt dynamics to be extremely driven by topography. We delineated the sub-basins using the hydrological tools implemented in the QGIS Geographic Information System Version 3.10, while we carried out the raster dataset processing through python scripts (https://www.python.org/). In those sub-basins where a meteorological station measuring snow height is present, we evaluated the aggregated data series. In detail, we converted the snow height into SWE values, attributing a snow density equal to 300 kg m^{-3}, as suggested in the literature about the properties of snow in the Italian Alps (Guyennon et al., 2019; Pistocchi, 2016). Then, we correlated the station daily values with the corresponding SWE data and calculated R^2 as an index of the fit.

To investigate the relationships between rainfall, snow melting and landslide triggering, we extracted the dated events for which both snow data and rainfall data were available at the reference meteorological station in the preceding 30 days (i.e. landslide events within an SWE sub-basin with a meteorological station). From the snow height data series, we obtained the daily snow accumulation (positive values) or snowmelt (negative values), which we converted into SWE. We then observed if snowmelt had occurred before the event, in which amounts and if it was concentrated in specific time windows. Based on these observations and the verified correlation between raster and station datasets, from the SWE dataset we derived three summary variables for the whole study area to be tested within the susceptibility model. They are: i) average number of melting events occurring over two or more 8-day periods in a hydrological year (ME_n); ii) average amount of melting recorded over four 8-day periods considering the whole data series (ME_{mean}); iii) maximum amount of melting recorded over four 8-day periods in the whole data series (ME_{max}).

3.1.3 Other variables

Besides climate variables, we inserted in the statistical model also classical relief, NDVI, land use and geological predictors. We derived the relief predictors (elevation, slope, aspect, curvature, flow accumulation) from the Aosta Valley Digital Elevation Model at 2 m × 2 m horizontal resolution (https://geoportale.regione.vda.it/mappe/informazioni-geoscientifiche), reclassified at 10 m × 10 m, using ESRI ArcGIS® 10.2.2 Spatial Analyst tools.

For land use, we reclassified the latest version of CORINE Land Cover (CLC2018, Büttner et al., 2017) into six classes, namely natural vegetation, agricultural, artificial, bare rock, glacier and water. For model training purposes, we considered possible land use changes at landslide triggering locations using the CLC change products and we attributed to the landslide cells the land use mapped on the initiation date. In terms of land use, we also derived two continuous maps based on the Normalized Difference Vegetation Index (NDVI). For the study area, we calculated maximum and average NDVI maps using LANDSAT 8 OLI/TIRS C1 Product Level 1 data, with a horizontal resolution of 30 m and a temporal resolution of 15 days, for the period 2014–2020. We excluded winter images (November–April) from the analysis because of snow cover.

For geology, we used the 1:10000 map available on the Aosta Valley geoportal (https://geoportale.regione.vda.it/mappe/informazioni-geoscientifiche). We reclassified the map into seven classes, namely gravitational deposits, glacial deposits, alluvial deposits, tectonized rock, hard rock, artificial soil and glacier and water bodies. Analysing shallow landslides occurring in soils, with this reclassification we tried to emphasize differences in loose soils (deposits) rather than based on lithology.

3.2 Susceptibility analysis

3.2.1 Generalized additive models

We developed the landslide susceptibility assessment using a Generalized Additive Model (GAM - Hastie and Tibshirani, 1990). A GAM is a data-driven approach merging characteristics of generalized linear models and additive models. Several authors successfully applied it for landslide susceptibility analyses (Bordoni et al., 2020a; Goetz et al., 2015; Knevels et al., 2020; Lombardo and Tanyas, 2020; Petschko et al., 2014). A generalized linear model relates the response variable to predictor variables through linear coefficients. In addition, a GAM can relate the response variable to a predictor variable through smooth functions. As for the linear terms, automatic fitting of the smooth functions is possible. Also, as in generalized linear models, it is possible to use link functions to model non-normal response variables. These characteristics make GAMs a powerful predictive tool with the advantage of an easy interpretation of the effect of each predictor variable on the model outcome (e.g., Brenning et al., 2015), which is a great advantage to investigate the physical meaning of the fitted model.

In detail, we applied the GAM to a binomial (1 – 0) response variable (landslide – non-landslide), thus deriving the probability of class 1 (π_1) from a logit link function:

$$logit(\pi_1) = \ln\left(\frac{\pi_1}{\pi_0}\right) = \ln\left(\frac{\pi_1}{1-\pi_1}\right) = \beta_0 + f_1(x_1) + f_2(x_2) + \cdots + f_i(x_i) \quad (4)$$

where π_0 is the probability of class 0, β_0 is the model intercept, x_i are the predictor variables and f_i their smoothing functions. We used the *gam* function as implemented in the *mgcv* library of R (Wood, 2020), automatically estimating the degree of smoothness of the predictor functions by restricted maximum likelihood. For the aspect variable, we fitted cyclic cubic regression splines to account for similarities between low and high values (0°–360°). For all the other variables, we fitted low rank thin plate regression splines. We performed model prediction on a 10 m × 10 m grid.

3.2.2 Model optimization

For landslide susceptibility estimation, we excluded from the investigation all those areas classified as bare rock, glacier and water in the land use map, and those classified as hard rock and glacier and water bodies in the geological map. We retained artificial areas because of possible landslides triggered along roads and those classified as tectonized rock because this tends to be weak rock with a mechanical behaviour similar to soils.

Since the binomial GAM model needs both landslide and non-landslide input data to train the model, we randomly sampled 300 non-landslide points with the only constraint of selecting them at least 100 m apart from each other and from the landslide points. We selected the number of points to match the 298 dated landslide events on an approximately 1:1 ratio as in several previous studies (e.g., Bordoni et al., 2020a; Chen et al., 2018). To optimize the model, we used all the dated landslides as input.

Firstly, we derived a model without climate-related variables to be used as reference. Then we tested several model configurations, using each time only one NDVI variable (mean or maximum) and only one pair of rainfall- and snow-related variables (nine combinations in total). Statistically, we optimized the model through penalization of the smooth functions (Wood, 2020). We did this by modifying the smoothing penalty with an additional shrinkage term (gamma = 1.5). This resulted in zero functions for unimportant variables, which were therefore removed from the model. Besides checking for statistical significance, we evaluated the form of the smooth functions. First, we allowed for function fitting without constraints. If we could explain their form physically, we kept the functions as they were; otherwise, we fixed the smoothing parameter or the number of basis functions. If also when constraining the behaviour of the variable we could not explain it, we removed the variable from the model. We implemented this approach especially for the climate variables. We derived them in such a way that the susceptibility should increase as their value increases. As an example, an increase in rainfall threshold exceedances (or in the number of snowmelt events) should lead to increasing stress cycles in a soil and therefore to an increase in susceptibility. Therefore, we could accept as physically significant for such variables only monotonous (increasing) smooth functions.

We considered a model optimal if it included both a rainfall- and a snowmelt-related variable and if the maximum concurvity between variables did not exceed 0.80. For both the reference and the optimal models, we evaluated the importance of the different variables through the decrease in deviance explained (Dev). We calculated this as in Knevels et al. (2020). Also, we used the decrease in deviance explained as a discriminant ($\geq$ 1%) to retain categorical variables in the model. In addition, for categorical variables we investigated the impact on the model outcome of different classes deriving odds ratios (Odds). Odds ratios represent the chance of an outcome, given a certain class in comparison to a reference one. For a class, odds greater than one indicate a higher chance of occurrence for the modelled outcome (in our case landslide) in comparison to the reference (Szumilas, 2010). We fixed as reference the most common class. In addition, we derived a contingency table with the threshold for landslide and non-landslide separation set at 0.5, and the area under the receiver operating characteristic (AUROC) curve derived from map areas (Chung and Fabbri, 2003) for both the dated and non-dated landslides. Finally, we reclassified the reference and the optimized maps 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), as in Samia et al. (2018), and we calculated the differences in terms of the class transition between the two maps.

3.2.3 Model validation

We validated the optimized model through a *k*-fold approach. We set up a validation framework with five folds and 20 repetitions, for a total of 100 fitted models each of which was trained with 80% of the points (dated landslides) and tested with the remaining 20%. For each model, we verified that the maximum concurvity between two variables did not exceed 0.80. To confirm robustness and consistency, we counted the number of times that each variable resulted in being statistically significant (not penalized to zero) and we checked the physical significance of their smooth functions. In addition, as with the model optimization, we calculated the deviance explained by the variables, contingency tables for both training and test points and AUROC values for training, test and non-dated landslide points.

4 Results

4.1 Model variables

4.1.1 Rainfall-related variable

Table 2 summarizes the association of landslides with meteorological stations recording hourly rainfall and specifies how many of them we could use for the exploratory analysis (first approach). Of the 45 landslide single dates theoretically covered by rainfall hourly data, we had to discard four complete records due to a temporary lack of data. In addition, for two landslide dates, cumulative rainfall was available only for up to 12 and 72 hours. From the analysis we conducted on these landslide dates, no specific rainfall duration was found to be critical. In terms of the percentage of landslide initiations associated with a rainfall event over the 75^{th} percentile of the corresponding monthly maxima distribution, we calculated 26%, 23%, 28%, 33%, 38%, 48%, 45%, 41%, 43% for 1-, 3-, 6-, 12-, 24-, 48-, 72-, 96-, and 120-hour periods, respectively. As an example, Fig. 4 shows these relationships for the seven landslides associated with ST04. In this case, two events show a cumulated rainfall above the corresponding monthly 75^{th} percentile for all durations and a third event just for 120 hours.

Table 2. Summary of the association between meteorological stations recording hourly rainfall, landslide and number of events above the rainfall threshold for landslide occurrence (EAT) recorded in Winter–Spring (Nov–Jun) and Summer–Autumn (Jul–Oct).

Meteo station	*Total landslides*	*Landslides covered by data*	*Landslide single dates*	*EAT Nov–Jun*	*EAT Jul–Oct*
ST01	17	3	3	15	17
ST02	7	3	2	13	10
ST03	42	7	7	17	5
ST04	10	7	7	14	16
ST05	1	0	0	10	10
ST07	1	0	0	5	14
ST08	0	0	0	20	16
ST09	6	2	2	6	13
ST10	56	0	0	7	14
ST11	1	0	0	10	21
ST12	6	3	3	4	3
ST13	54	2	2	6	4
ST14	57	7	5	13	4
ST15-16	19	5	4	13	5
ST17	21	19	10	9	5
TOT	298	58	45	162	157

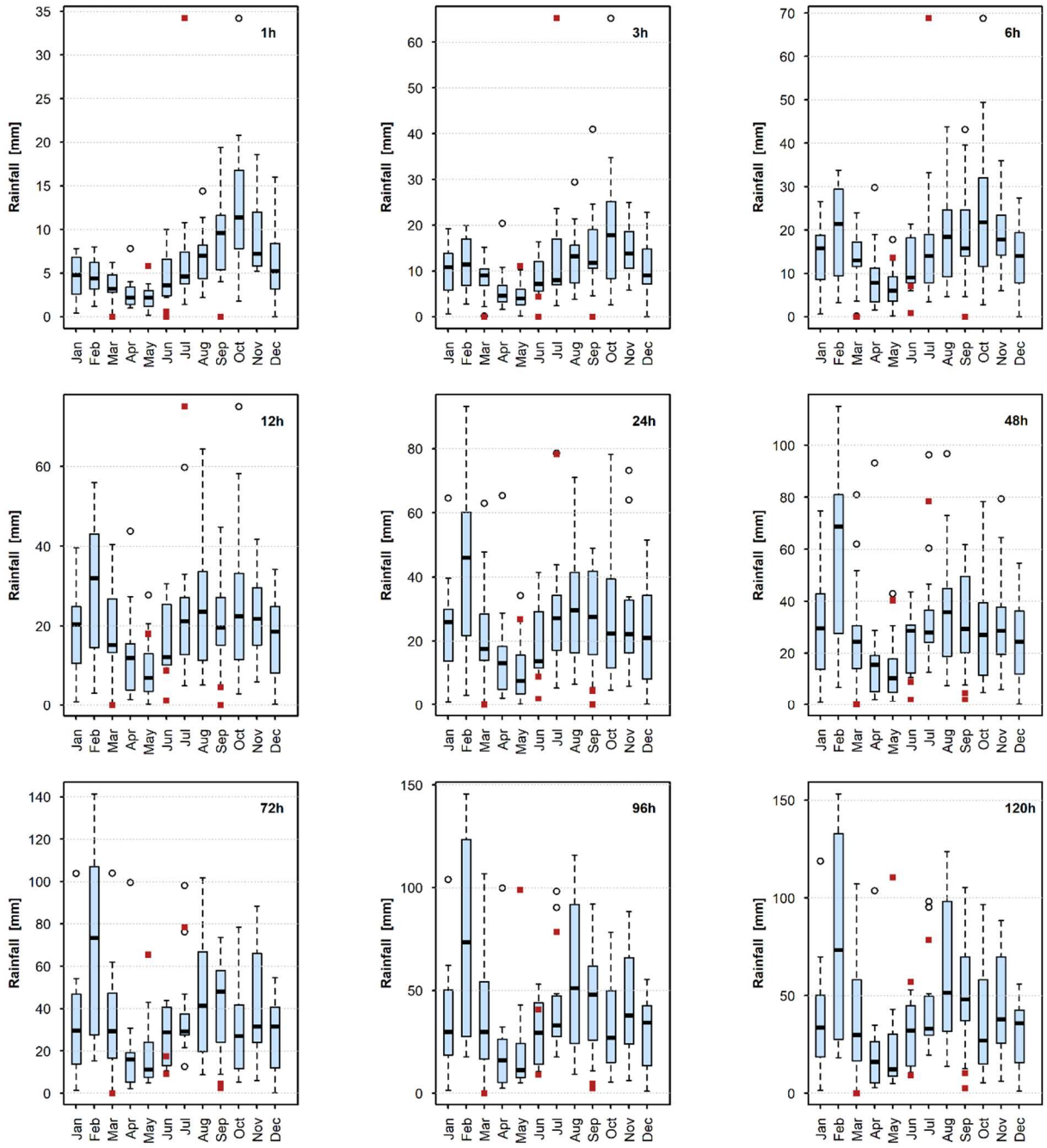


Fig. 4. Boxplots showing the distribution of monthly maxima for different rainfall durations (top-right corner of each panel) and cumulated rainfall preceding landslide initiation (filled squares). For the boxplot, the *boxes* represent the interquartile range (IQR), the line inside the box is the median, whiskers extend 1.5 IQRs, and the open circles are the outliers.

To proceed with the second approach, we adjusted the threshold proposed by Aleotti (2004) for Piedmont. Based on the available data, we could link only three landslide dates to rainfall conditions exceeding the threshold proposed by Aleotti (2004). In addition, two of these dates belonged to the same storm event that affected Aosta Valley in mid-October 2000. Therefore, to have at least three events other than the October 2000 flood above the threshold, we lowered it as shown in Fig. 5a. The total number of available landslide dates above the threshold could appear low (five dates equals 11% of the total) but some bias could be present because of local rainfall variability (distance between the location of the meteorological station and the landslide). In addition, the threshold must represent

extreme conditions. Based on this newly defined threshold, we derived the events above the threshold variables (EATs) that are shown in Fig. 5.

The climate characterizing the study area can be divided into two broad seasons. The first, Winter–Spring (November–June) is characterized by persistent unstable weather and temperatures that at different altitudes move across zero, possibly causing snowmelt events. The second, Summer–Autumn (July–October) is mainly characterized by thunderstorms and convective events of shorter duration. Table 2 shows local differences in the seasonal frequency of events above the threshold (e.g., at ST07 and ST08, which are nearby and at similar altitudes); however, the events tend to balance over the year. We believe that this balance is an indication that the EAT variables are able to capture events with characteristics typical of different seasons and that therefore they can be considered good descriptors of the average annual tendency of a particular area to be affected by intense rainfall.

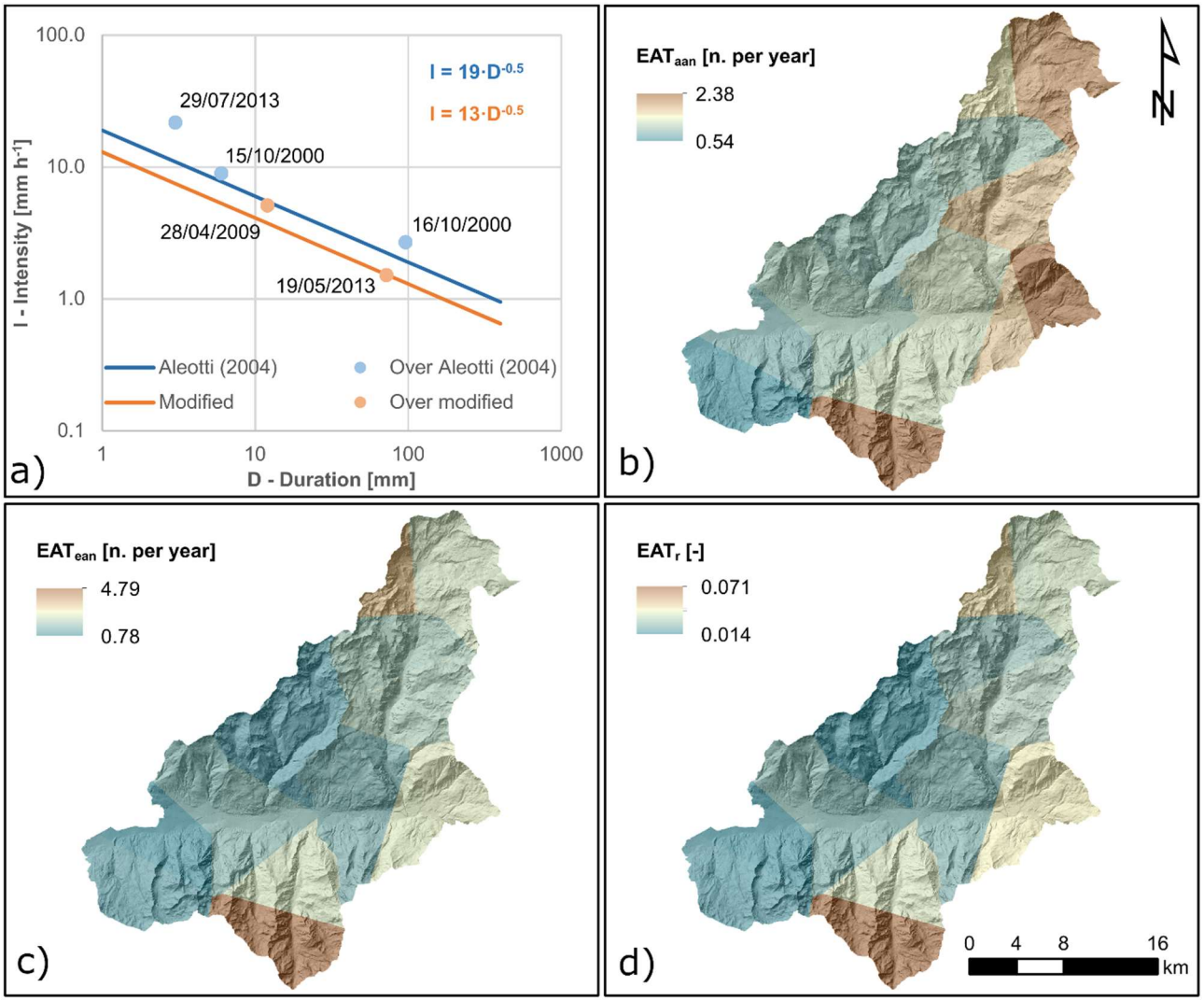


Fig. 5. a) Comparison between the rainfall-threshold proposed by Aleotti (2004) and the one modified for this study, b) average annual events above threshold (EAT_{aan}), c) effective average annual events above threshold (EAT_{ean}), and d) above threshold event ratio (EAT_r).

4.1.2 Snow-related variable

The correlation between the aggregated SWE data and the time series recorded at the seven meteorological stations with a snow sensor showed a moderate to good agreement. At two stations, the R^2 coefficients were found to be higher than 0.6, at three stations they showed values around 0.4 and at the remaining two stations, Nus Saint-Berthelemy (ST14) and Saint-Christophe-Aeroporto (ST17), they scored below 0.2. These low values are due to many station values reporting an absence of water stored as snow, while the area average SWE data series reported up to 50 cm. Fig. 6 shows two examples of the correlations. Overall, we considered the aggregated data to be representative of the snow dynamics and suitable for further analyses.

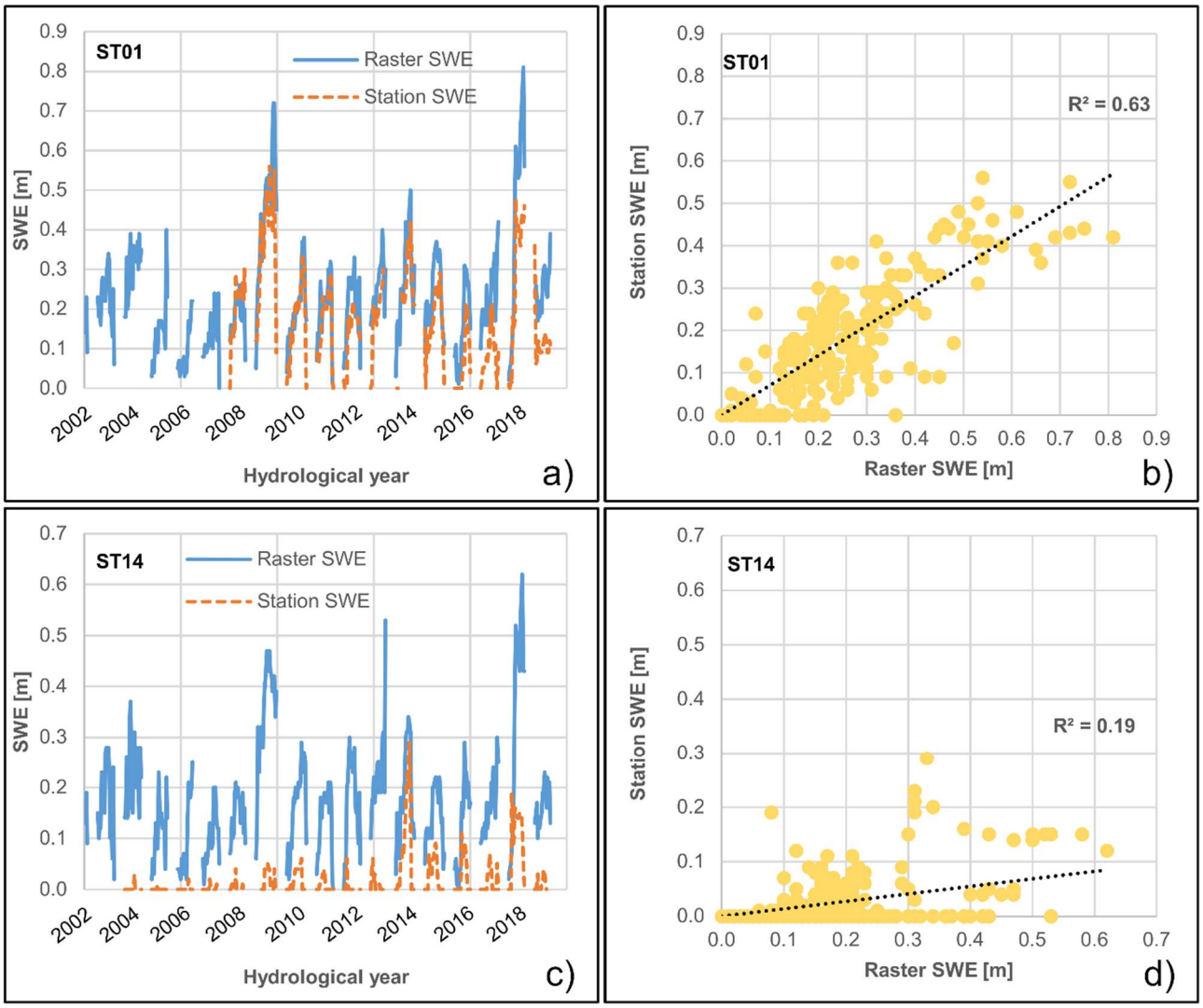


Fig. 6. Comparison and correlation between the time series of Snow Water Equivalent (SWE) obtained at stations and from the aggregation of the raster dataset over sub-basin polygons.

We performed the analysis of rainfall and snowmelt dynamics in the days preceding the landslide initiation for 16 events, three of which occurred on the same date in the same area (i.e., 14 independent landslide dates). In 50% of the cases no snowmelt occurred. In the other half of the cases, the cumulated snowmelt after 30 days ranged from 1.9 to 34.8 mm. In only two cases, we observed snowmelt occurring in the week preceding the event. In both cases, snowmelt also increased, although

discontinuously, extending the observational time window. Fig. 7 shows two typical examples, one characterized by discontinuous snowmelt since the day immediately preceding the event (Fig. 7a) and the other characterized by snowmelt only beyond the week preceding the event (Fig. 7b). Although the available data are limited, we concluded that snowmelt can influence landslide initiation, as a preparatory factor, through a prolonged (15 to 30 days) or repeated action that favours high volumetric water contents in soils (Subramanian et al., 2020). For this reason, we derived ME_n, ME_{mean} and ME_{max} as variables to be tested in the susceptibility analysis (Fig. 8).

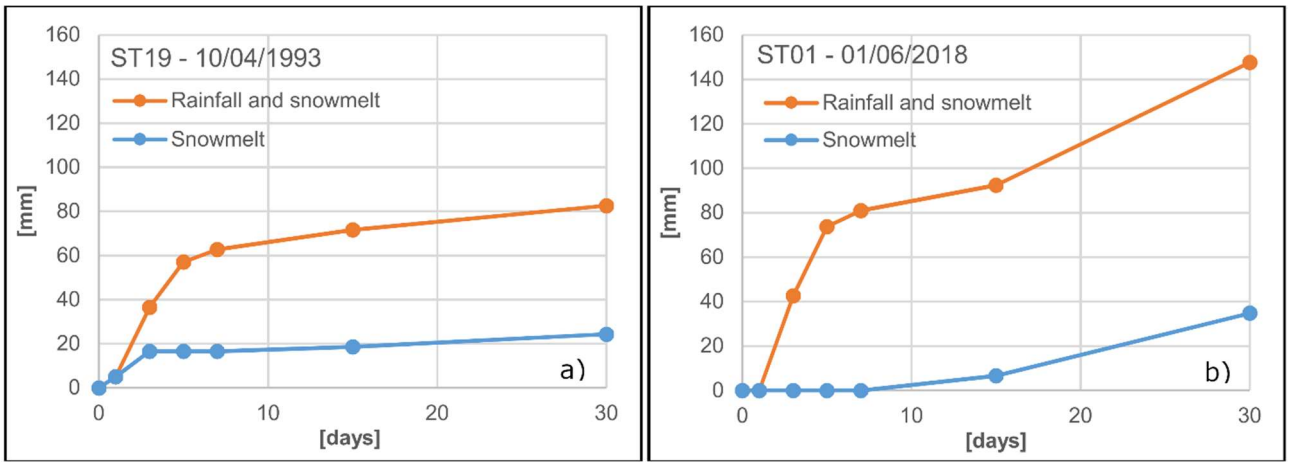


Fig. 7. Rainfall and snowmelt dynamics in the 30 days preceding two landslide events.

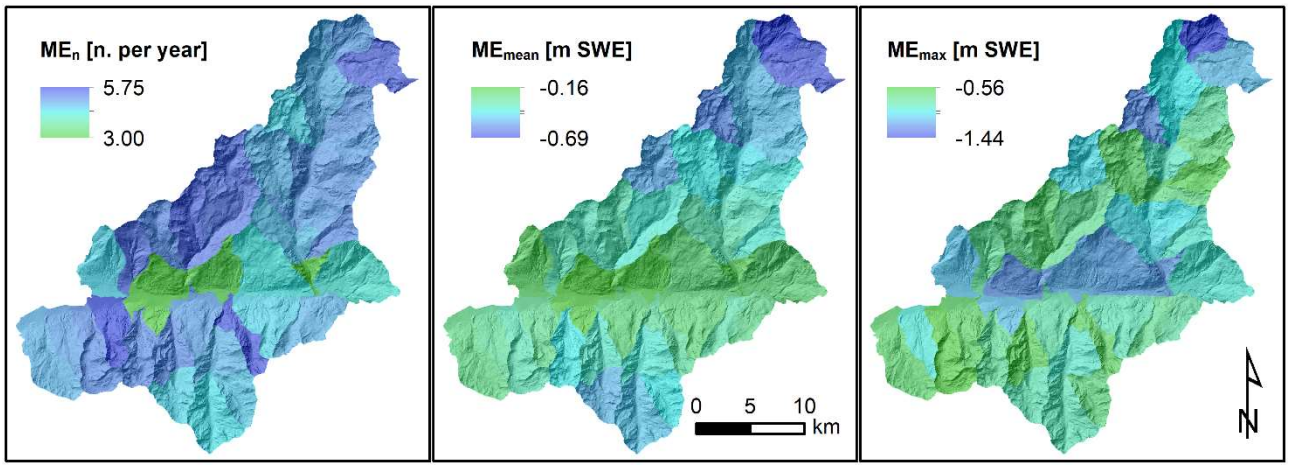


Fig. 8. Derived maps of a) average number of melting events occurring over 16-day period in a hydrological year (ME_n), b) average melting amount recorded over 32-day periods considering the whole data series (ME_{mean}), c) maximum melting amount recorded over 32-day periods in the whole data series (ME_{max}). Since snowmelt is a depletion of the SWE storage, we considered the values negative.

4.1.3 Other variables

Fig. 9 shows the maps of the relief variables, the NDVI products, and the reclassified CORINE Land Cover and geology. In addition, we transformed the flow accumulation variable into logarithmic values and we tested both the original and transformed predictor in the statistical model.

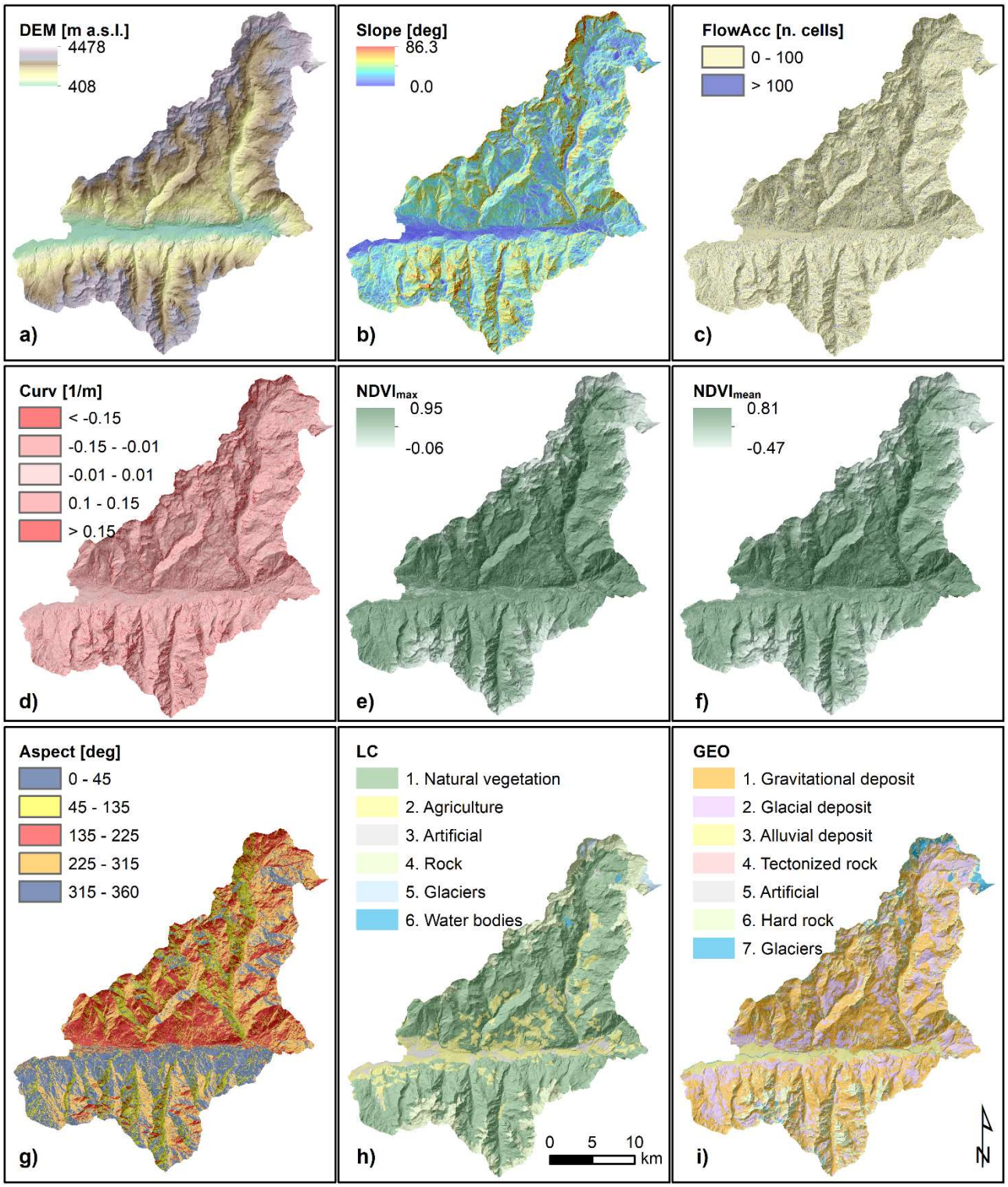


Fig. 9. Maps of the variables used as input in the susceptibility analysis, a) elevation (DEM), b) slope angle (Slope), c) flow accumulation (FlowAcc), d) curvature (Curv), e) maximum NDVI value ($NDVI_{max}$), f) mean NDVI value ($NDVI_{mean}$), g) orientation of the downward sloping terrain (Aspect), h) reclassified land cover (LC), i) reclassified geology (GEO).

4.2 Susceptibility analysis

4.2.1 Model optimization

Figure 10 shows the modelling area, with the trivial zones removed, and the points (dated landslides, non-dated landslides, non-landslides) used for model training and evaluation. The selection of variables through penalization for the definition of the reference model retained the elevation, aspect, slope angle, curvature and land cover. The highest concurvity, attributed to the parametric component of the model, was 0.62.

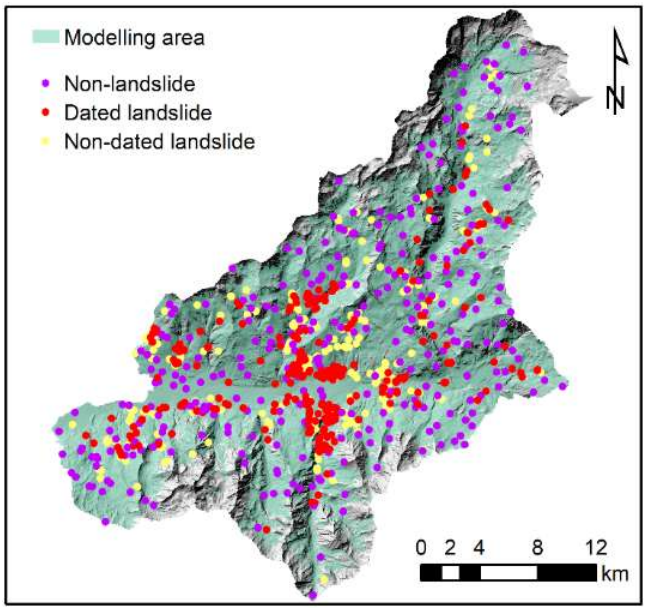


Fig. 10. Modelling area and points used for model training and evaluation.

Table 3 summarizes the model optimizations obtained including different pairs of rainfall and snowmelt variables. For all variable combinations, the penalization excluded the flow accumulation (normal or logarithmic values) and the NDVI (mean or maximum) variables used. The only combination of variables resulting in an optimized model that includes both rainfall and snowmelt was the one with EAT_{ean} and ME_n (Model ID 4 in Table 3). For this model, the highest concurvity was 0.63. In addition, the optimized model kept all the variables included in the reference model. In terms of the physical validity of the smooth functions, we had to put constraints for three variables, namely the slope angle (also in the reference model), EAT_{ean} and ME_n. Fig. 11 shows both the unconstrained and constrained smooth functions. For the slope, we cut off the values at 55° to reduce the width of the confidence intervals and we set the maximum number of basis functions to three, to avoid local minima and maxima that we could not explain. The constrained smooth function shows a maximum between 35° and 40°, with an almost constant (slightly inverse) relationship with the susceptibility to landslide occurrence for higher values. We interpreted this as the transition to shallower and coarser soils and therefore to a different, predominant type of landslide (e.g., debris flow; Ponziani et al., 2020). For EAT_{ean}, we set the maximum number of basis functions to three to guarantee a monotonous increasing behaviour of the variable. For ME_n, we achieved the same result by fixing the smoothing parameter to five. For us, increasing exceedances of the rainfall threshold and of the number of melting events correspond to increasing stress cycles in soils and therefore to

an increase in susceptibility. Like the slope angle, we attributed the monotonous decreasing relationship between elevation and susceptibility to variable soil properties. Besides, we did not exclude that the form of the smooth function could be related, to some extent, to a possible bias in the inventory, that might include more events close to residential areas and infrastructures (i.e. at low elevations) rather than in poorly populated valleys (i.e., at high elevations). Regarding the aspect, due to solar radiation the slopes facing south are the most exposed to diurnal cycles of temperature and therefore experience the highest rates of evapotranspiration in summer and (possibly) snowmelt in winter, deeply affecting the soil water content. We considered the form of the smooth function, with a maximum around 180°, as representative of these processes. Curvature indicates an increasing susceptibility for concave slopes.

Table 3. Summary of model optimizations including different pairs of rainfall and snowmelt variables. For variable acronyms, refer to Fig. 5, Fig. 8 and Fig. 9.

Model ID	*Rainfall variable*	*Snowmelt variable*	*Raifall stat. sign.*	*Snowmelt stat. sign.*	*Rainfall phys. plaus.*	*Snowmelt phys. plaus.*	*Other variables stat. sign. and phys. plaus.*
1	EAT_{aan}	ME_{n}	No	Yes	No	Yes	DEM, Aspect, Slope, Curv
2	EAT_{aan}	ME_{mean}	No	Yes	No	Yes	DEM, Aspect, Slope, Curv, LC, GEO
3	EAT_{aan}	ME_{max}	No	Yes	No	No	DEM, Aspect, Slope, Curv, LC
4	**EAT_{ean}**	**ME_{n}**	**Yes**	**Yes**	**Yes**	**Yes**	**DEM, Aspect, Slope, Curv, LC, GEO**
5	EAT_{ean}	ME_{mean}	No	Yes	No	Yes	DEM, Aspect, Slope, Curv, LC, GEO
6	EAT_{ean}	ME_{max}	Yes	Yes	Yes	No	DEM, Aspect, Slope, Curv, LC
7	EAT_{r}	ME_{n}	No	Yes	No	Yes	DEM, Aspect, Slope, Curv
8	EAT_{r}	ME_{mean}	No	Yes	No	Yes	DEM, Aspect, Slope, Curv, GEO
9	EAT_{r}	ME_{max}	No	Yes	No	No	DEM, Aspect, Slope, Curv

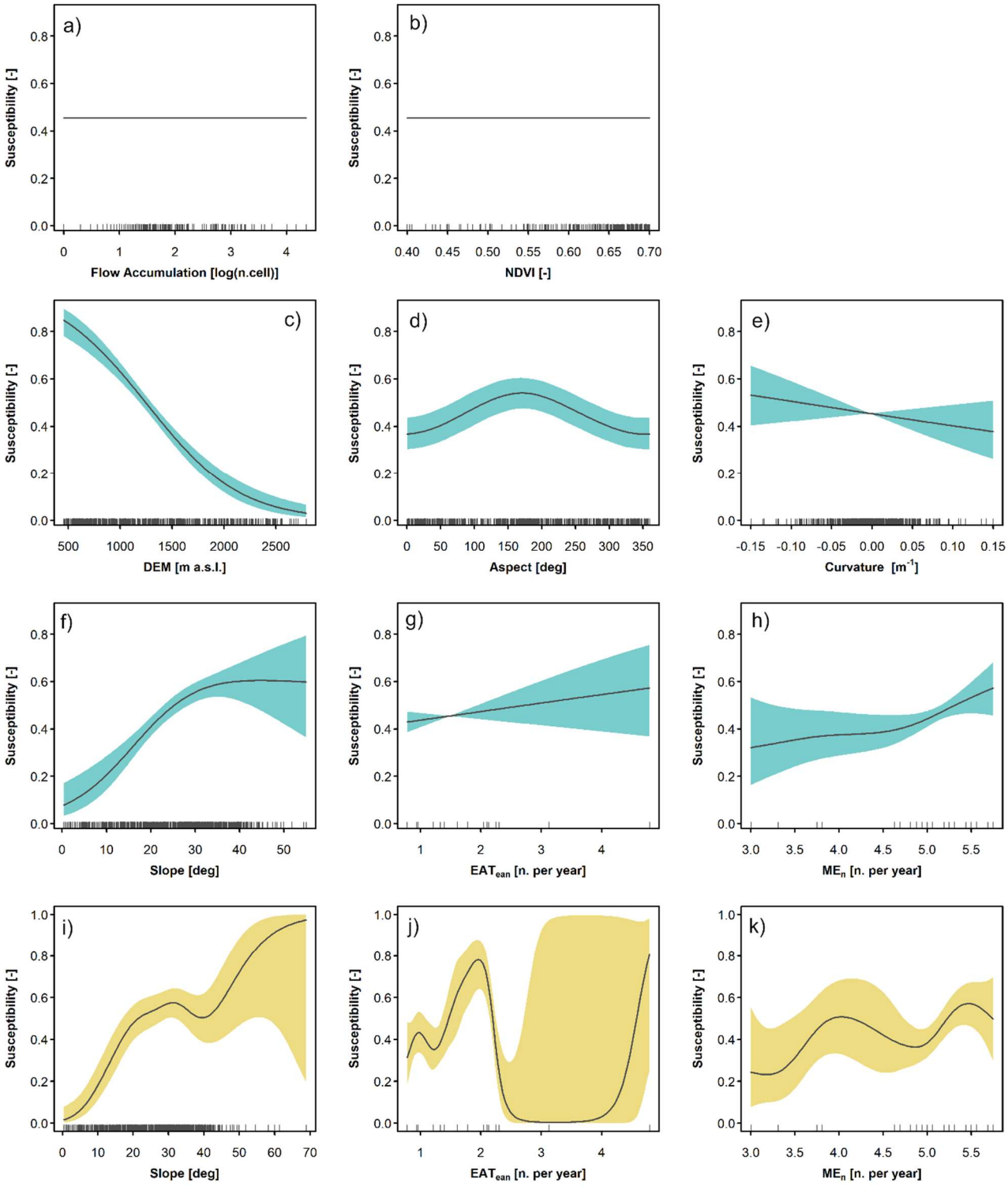


Fig. 11. Smooth functions of the variables used as inputs in the optimized model with their confidence interval, a) flow accumulation, b) maximum NDVI value, c) elevation (DEM), d) orientation of the terrain (Aspect), e) Curvature, f) slope angle (Slope), g) effective average annual events above threshold (EAT_{ean}), h) average number of melting events occurring in a hydrological year (ME_n); and non-constrained smooth functions, not included in the model, of i) Slope, j) EAT_{ean} and k) ME_n.

Table 4 and Fig. 12 show the comparison and the differences between the reference and the optimized model. In both models, the relief variables (DEM, Slope, Aspect) explain the majority of the variability (95% and 91% in the reference and in the optimized model, respectively). Since gravity is

the driving force of landslides, we considered the result reasonable. However, in the optimized model the climate variables overall explain 5% of the variability, which is higher than the sum of land cover and geology, variables very often included in such analyses. Among the two, snowmelt appears to be more important than rainfall.

The signal of the snowmelt events within the model can be considered seasonal. Fig. 13 shows that in Winter–Spring (November–June) two out of three (68%) landslides occur in areas characterized by ME_n values larger than 5, while in the Summer–Autumn period (July–October) only half (53%) of the landslides initiate in the same areas. Also, almost all the Winter–Spring landslides (92%) present ME_n values larger than 4.5, while only 72% of the Summer–Autumn landslides are characterized by the same values. However, the landslide inventory used as input is temporally biased due to the extreme event of October 2000, as 220 out of the 244 dated landslides in Summer–Autumn are related to this event, whereas only 54 landslides occurred in Winter–Spring. Therefore, the form of the smooth function related to ME_n is most probably related to the seasonality shown, but the deviance explained remains rather low because of this difference in the total number of landslides in the two seasons. Regarding EAT, since the variable is able to include different types of intense weather events (both short and prolonged, Table 2), the very low deviance explained is probably related to its poor spatial resolution rather than to seasonality.

The addition of the climate variables also caused an improvement in the modelling of the landslide points, evidenced by an increase of the true positive rate by 2.4% and of the AUROC of the dated landslides by 0.8%. The similar AUROC values returned by the two models for the non-dated landslides suggest the occurrence of these events in different topographical conditions rather than under different climate conditions, in comparison to the dated events.

Table 4. Summary of the reference and optimized model characteristics used for evaluation and comparison. See Fig. 9 for variable abbreviations.

Evaluation parameter	***Reference Model***	***Optimized Model***
Deviance explained DEM	68%	66%
Deviance explained Slope	24%	20%
Deviance explained Aspect	3%	5%
Deviance explained FlowAcc	0%	0%
Deviance explained Curv	0%	0%
Deviance explained NDVI	0%	0%
Deviance explained EAT	NA	1%
Deviance explained ME	NA	4%
Deviance explained LC	1%	1%
Deviance explained GEO	4%	3%
Odds LC2	1.58	1.54
Odds LC3	0.59	0.54
Odds GEO2	1.23	1.20
Odds GEO3	0.53	0.49
Odds GEO4	0.00	0.00
Odds GEO5	0.00	0.00
True Positives	74.8%	77.2%
True Negatives	71.7%	73.0%
Concurvity	0.62	0.63

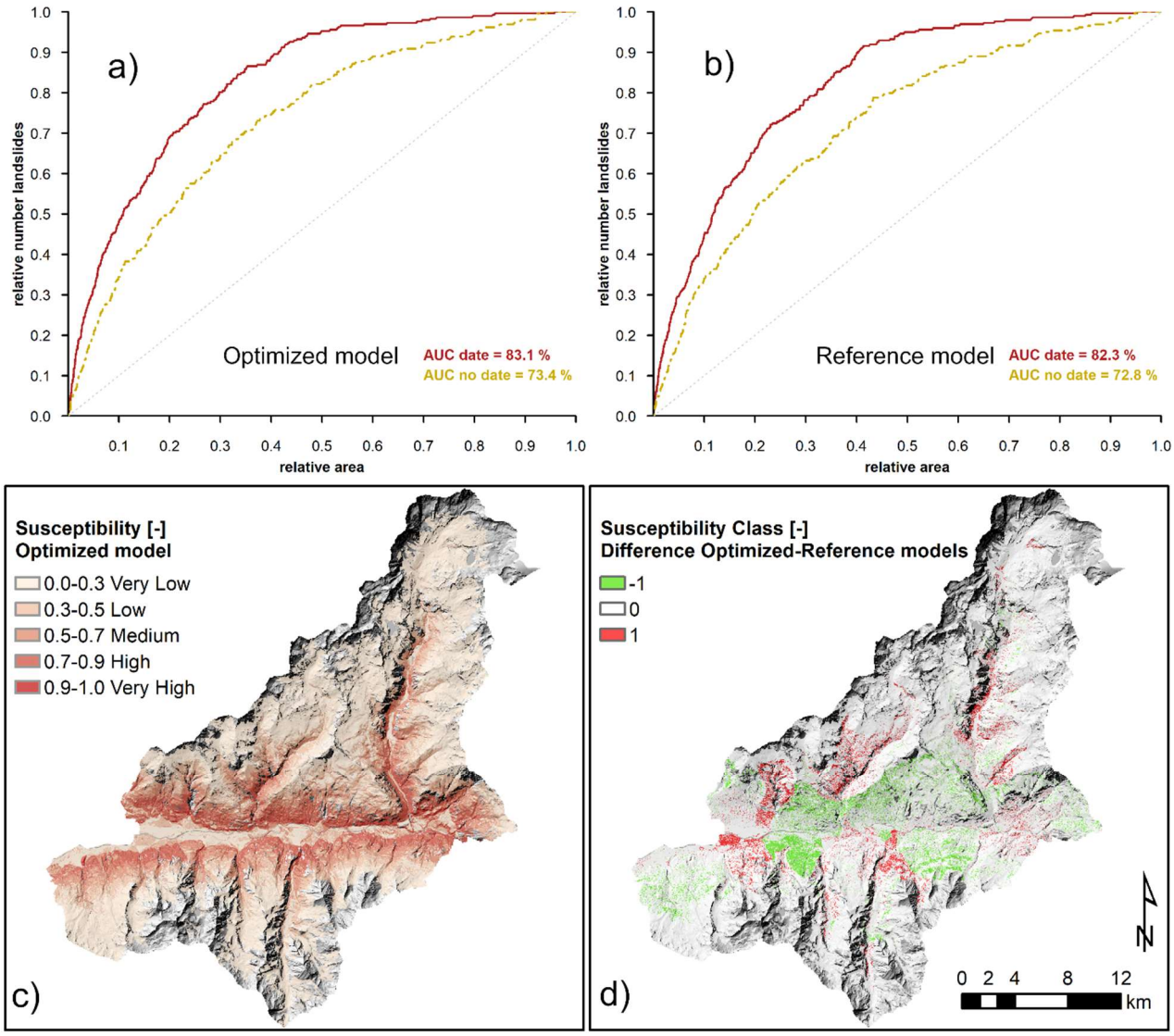


Fig. 12. ROC curves with AUC values for the a) optimized and b) reference models; maps showing c) the susceptibility to shallow landslides obtained from the optimized model and d) the differences with the map obtained from the reference model.

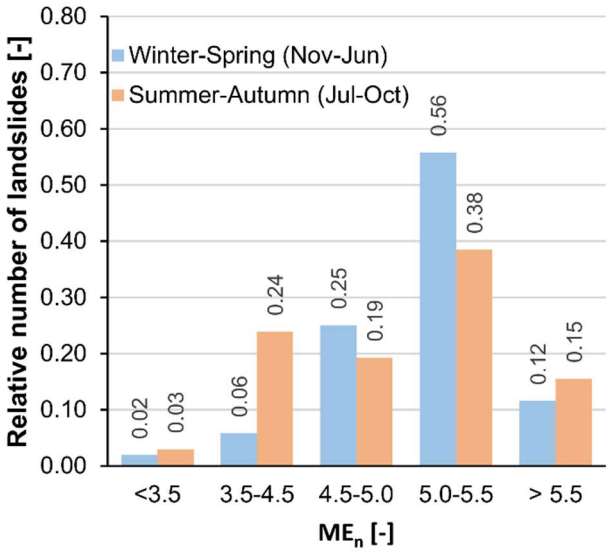
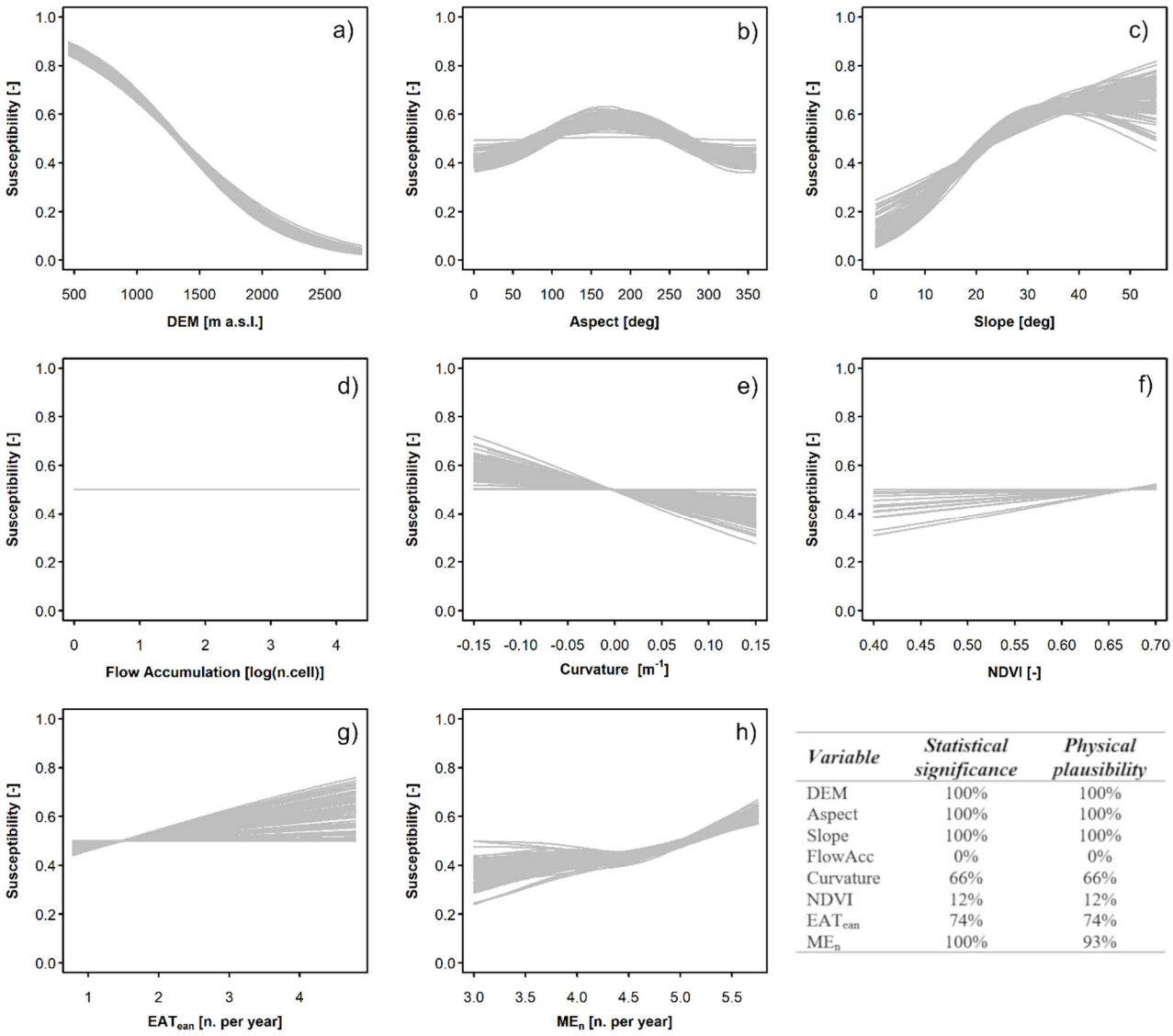


Fig. 13. Relative number of landslides occurring in areas characterized by a different average number of melting events (MEn) in Winter–Spring and in Summer–Autumn.

The map obtained from the optimized model (Fig. 12c) shows that the areas most susceptible to shallow landslide susceptibility are located at low elevations along the axes of the major valleys. In terms of differences with the reference model, the introduction of the climate variables affected 11.0% of the modelling area. In detail, 6.3% of the area passed to a less susceptibility class, while 4.7% experienced an increase in susceptibility. Transitions only occurred between contiguous classes. The slopes on both sides of the main E–W valley are the areas most affected by the transitions. In addition, transitions towards higher susceptibility classes occurred along the Valtournenche valley bottom and in the valley north of Nus.

4.2.2 Model validation

Fig. 14 shows the results of the *k*-fold validation in terms of fitted smooth functions. It also presents a summary of their statistical and physical significance, as well as the variability of different variable importance and model performance indices.



Variable	*Statistical significance*	*Physical plausibility*
DEM	100%	100%
Aspect	100%	100%
Slope	100%	100%
FlowAcc	0%	0%
Curvature	66%	66%
NDVI	12%	12%
EAT_{ean}	74%	74%
ME_n	100%	93%

Fig. 14. Smooth functions fitted in the 100 models of the *k*-fold validation for the continuous variables used as model input and summary of their statistical and physical significance: a) elevation (DEM), b) orientation of the terrain (Aspect), c) slope angle (Slope), d) flow accumulation, e) curvature, f) maximum NDVI value, g) effective average annual events above threshold (EAT_{ean}), h) average number of melting events occurring in a hydrological year (ME_n).

In terms of statistical significance (Fig. 14), DEM, Slope and Aspect proved to be the most consistent variables, never being penalized to zero functions. The snowmelt-related variable (ME_n) behaved similarly; however, we fixed its smoothing parameter as in the optimized model. The rainfall-related variable (EAT_{ean}) and curvature showed similar behaviour. They were penalized to zero functions in around 30% of the cases, which is partly attributable to the lower importance in the model (lower decrease in deviance explained, see Figure 15) and partly to the large confidence intervals especially around their minimum and maximum values. Especially for EAT_{ean}, the loss of structure of the smooth function could be related to the few training points (five) associated with the variable maximum value. Therefore, in a *k*-fold validation set up with five folds, it is inevitable that in some repetitions these areas are under-represented such that the EAT_{ean} function loses robustness. The area is one of the most remote of the entire study area and the lack of more landslide points most probably derives from an inventory bias.

In terms of the variables' importance (Fig. 15a), the results of the *k*-fold validation show little variability in comparison to the optimized model fitted with the full dataset. In terms of decrease of deviance explained, DEM, Slope and Aspect remain the most important, explaining together on average 91% of the model variability (minimum 83%, maximum 98%). The climate variables together presented an average decrease of deviance explained of 4% (minimum 0%, maximum 11%). Also, in 90% of the cases, excluding ME_n from the model resulted in a greater decrease of deviance explained than excluding EAT_{ean} (ranges 0–8% and 0–3%, respectively). When characterized by a non-zero function, curvature is usually the variable leading to the lowest decrease of deviance explained when excluded from the model. Looking at the odds ratios (Fig. 15b), consistently with the optimized model, the agriculture land use (LC2) seems to have a greater chance than natural vegetation (the reference LC1) of experiencing landslides, while the presence of artificial soils (LC3) shows an odds variability between 0.2 and 1.0. However, the artificial soils class is significant (*p-value* < 0.10) in only 4 out of the 100 fitted models (37 out of 100 for LC2), explaining the odds values tending to one and indicating no relationship with the investigated phenomenon. In terms of geological classes, alluvial deposits (GEO3) appear less prone than gravitational deposits (reference GEO1) to experience shallow landslides (odds < 1 and *p-value* < 0.10 in 35 out of 100 models), while there is no clear tendency for glacial deposits (GEO2, boxplot across one and *p-value* < 0.1 in a single case). Tectonized rocks and artificial soils (GEO4 and GEO5, respectively) show odds ratios equal to zero, meaning that no landslides are expected with these conditions. These classes never have a *p-value* < 0.1, probably because they occupy a small area and present very few training points. The maximum model concurvity is always lower than 0.7 (range 0.61–0.68) and it usually occurs between DEM and the model parametric components (LC and GEO).

Regarding the model performance (Fig. 15c), the true positives (TP) and true negatives (TN) rates calculated on the test sets are those presenting the widest interquartile range (IQR). However, the overall IQR values show a rather limited variability. They are between 66.7% and 73.3% for TN and between 71.2% and 78.3% for TP. For the training points, the width of the IQR is much reduced and the median value of TP is only slightly lower (76.5%) than for the optimized model (77.2%). Almost no variability is present in terms of AUROC values, which are included in the intervals 81.9–83.4%

and 81.3–83.3% for training and testing points (dated landslides), respectively. For non-dated landslides, AUROC ranges between 71.1% and 74.2%.

Overall, we considered that the *k*-fold validation confirmed the structure and the capabilities of the optimized model, from both a statistical and a predictive point of view. Regarding the climate variables, the validation confirmed their importance within the model, although it pointed out small criticalities that are probably attributable to landslide inventory biases in remote areas.

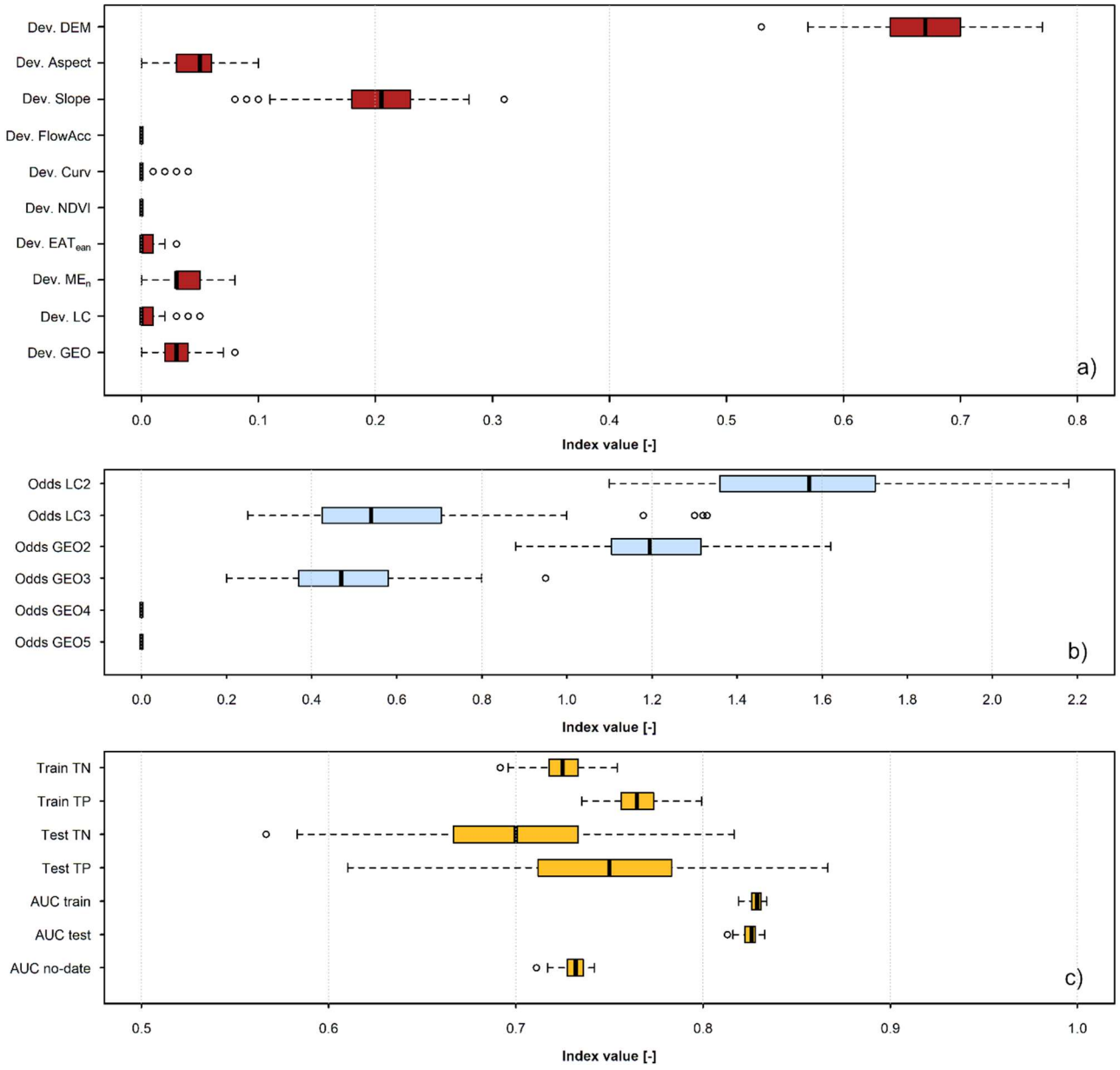


Fig. 15. Boxplots showing the variability observed during the *k*-fold validation regarding a) the decrease in deviance explained (Dev.) associated with model variables; b) the odds ratios (Odds) of land cover and geology classes (see Fig. 9 for variable abbreviations); c) true negatives (TN) and true positives (TP) rates for both training and testing sets and area under the ROC curve (AUC) values for training, testing and non-dated landslide points.

5 Discussion

In event-based landslide susceptibility assessments, several authors (Gassner et al., 2015; Knevels et al., 2020) recognized cumulated rainfall in a defined period of time as an important predictor in their statistical models. In this multi-temporal susceptibility analysis, we performed an exploratory analysis to link landslide initiation to rainfall characteristics, but we could not recognize a specific duration leading to landslide triggering. Therefore, we opted to derive a rainfall summary variable based on intensity–duration thresholds for landslide occurrence (e.g., Segoni et al., 2018a). In previous studies, authors (Segoni et al., 2018b, 2015) used such thresholds in combination with susceptibility maps to derive hazard assessments, but they never introduced them within the statistical susceptibility model. Dikshit et al. (2020) derived intensity–duration thresholds for landslide occurrence for their study area and calculated their annual exceedance probability, in a similar way as we did in this study. However, they used it to derive the temporal probability of landslide occurrence to be used alongside the susceptibility map for preliminary risk management. According to our findings, the use of the threshold exceedance variable within a susceptibility model represents a concise and versatile way to introduce a temporal (non-stationary), climatological factor in a multi-temporal (i.e., multi-event) susceptibility analysis. It is concise because it is just one variable and it is versatile because it can be rather simply calculated for different reference periods and scenarios. In addition, the output is a single product (a map) with a specific reference period, which can therefore be used for mid- to long-term planning.

Another innovative approach that uses rainfall variables for the definition of a non-stationary susceptibility is the one proposed by Bordoni et al. (2020b). The authors combined a spatial landslide susceptibility with a temporal landslide susceptibility, both derived from statistical models, by simple multiplication of the outputs. They derived the temporal susceptibility based on short-term cumulated rainfall (three days), antecedent rainfall (30 days) and soil moisture, thus being able to update it on a daily basis and through the evolution of a rainfall event. If our approach mainly answers the need for a planning instrument, the approach of Bordoni et al. (2020b) can be adapted for use as an operational early warning system. The two approaches can therefore be seen as complementary.

In terms of the importance of the variables within the optimized model, several authors report the great relevance of relief predictors (e.g., Petschko et al., 2014; van Westen et al., 2008), which is also evident from the high percentage of studies including these in the analysis (Reichenbach et al., 2018). Another variable that previous studies suggest as relevant is land use (e.g. Bordoni et al., 2020b; Glade, 2003; Knevels et al., 2020; Reichenbach et al., 2014). The moderate importance that it shows in this study is probably due to the broad land cover differentiation that we adopted. However, consistently with the previously cited studies, agricultural (i.e. bare or scarcely vegetated) soils proved more prone to the occurrence of shallow landslides than natural (i.e. mid to densely vegetated) soils. In terms of climate variables, the rainfall-related predictor (EAT_{ean}) showed an importance similar to that of the three-hour maximum intensity in the models developed by Knevels et al. (2020). The snowmelt-related predictor (ME_n) usually showed a higher importance than EAT_{ean}. This is probably due to the better spatial resolution of the former in comparison to the latter. This highlights the increasing need and demand for novel high-resolution (hourly) gridded rainfall datasets and adjusted radar products (e.g., Panziera et al., 2018).

In a recent review on shallow erosion processes over the Alps (including shallow landslides), Geitner et al. (2021) stated that the effects of a changing land use and a changing climate are often interrelated and difficult to separate. In general terms, they observe that land abandonment (including grasslands used as pasture) can lead to increasing instabilities until vegetation succession counterbalances the enlargement of unstable areas and that an increasing frequency of intense rainfall events as well as increasing frequency and intensity of snowmelt can increase soil failures. According to Ponti et al. (2021), for alpine and nival belts (i.e., above 2400 m a.s.l.) ground heating has a key role in controlling surface displacement and instability. Ground heating controls the snowmelt frequency and intensity as well as influencing permafrost degradation. Grassland vegetation species that are not able to adapt to soil displacement respond by moving towards higher elevations. Therefore, the variables included in the statistical model seem a good selection to map future changes in shallow landslide susceptibility. A more detailed variable for land use and vegetation could possibly be tested in future studies.

Steger et al. (2016) suggested that high performance indices in the evaluation of statistical models are not a guarantee of the geomorphic plausibility of the derived landslide susceptibility maps. To overcome the issue, they focused their attention on recognizing inventory biases and artifacts in the output map. By supervising the developed model, we intended to reach the same objective with a different approach. GAMs proved to be an optimal instrument since they allowed the behaviour of the predictors to be constrained without coercing them into an exact relationship with the dependent variable.

Regarding future perspectives, to actually make the map we derived non-stationary, EAT_{ean} and ME_n variables should be defined for future reference periods. Considering the biases that usually affect regional climate models (RCMs) in mountain environments in terms of rainfall and snow (Kotlarski et al., 2014; Matiu et al., 2020), the simple use of these model data could lead to unsatisfactory results. In addition, RCMs data may not have the necessary spatial resolution to capture and describe the processes investigated. Therefore, various options including high-resolution dynamical downscaling (e.g., Shou and Lin, 2020), statistical downscaling (e.g., Camera et al., 2017a) or simple bias correction (e.g., Lazoglou et al., 2020) should be evaluated to find the best trade-off in terms of output quality and computational effort. A dedicated study is therefore necessary.

For ME_n we showed a seasonal signal within the model that was partly smoothed out by a temporal bias in the inventory. Given that similar characteristics of the explanatory variables are necessary in the training and test sets and that they must cover the whole range of properties within the study area (Guzzetti et al., 2006), the low number of landslide occurrences in the Winter–Spring season (54) did not allow the training and testing of a seasonal model to further investigate the role of the climate variables in specific periods of the year. However, because of the different changes expected in the precipitation patterns in the different seasons over the Alpine area (Ban et al., 2020), the development of seasonal landslide susceptibility models for the detection of climate change impacts would be an extremely valuable contribution for future studies.

6 Conclusion

We presented an approach to derive and include climate variables, strictly related to the occurrence of shallow landslides, within a susceptibility analysis. The objective was to give a process-related

non-stationary configuration to the landslide susceptibility model to make it adaptable to the study of future climate changes. In detail, we focused on the evaluation of the statistical significance and physical plausibility of intense rainfall and snowmelt variables within a generalized additive model (GAM) and a *k*-fold cross-validation. We also evaluated the added value induced by their inclusion in the susceptibility analysis, considering as reference a model constructed with relief, land use and geological predictors.

We included in the susceptibility model an intense rainfall variable based on the frequency of exceedance of an intensity–duration threshold for landslide occurrence (effective average annual events above threshold - EAT_{ean}). The advantages of this variable are its conciseness (just one variable) and versatility, since it can be simply calculated for different reference periods and scenarios. Regarding snowmelt, we concluded that it influences landslide initiation through a prolonged (15 to 30 days) and repeated action (melting events in a hydrological year), therefore from a raster SWE dataset we derived and introduced in the model a variable counting the average number of melting events occurring over 16-day periods in a hydrological year (ME_n). The variable is particularly interesting because it can reproduce the dynamics of high-altitude environments, which are usually disregarded.

The statistical model optimization and *k*-fold validation confirmed the relevance of the climate predictors in terms of both statistical significance and physical plausibility. Within the model, elevation and slope angle explained most of the deviance, while the climate variables together accounted for just 5% of it (up to 11% in singular cases in the *k*-fold validation). This relatively low contribution is probably related to a seasonality of the snowmelt variable masked by a temporally-biased landslide inventory (220 landslides out of 298 occurred in October 2000) and to a low spatial resolution of the rainfall variable. However, the two variables caused an improvement in the model performance, in comparison to the reference model, of 2.4% in terms of the true positives rate and 0.8% in terms of the AUROC. In doing so, the susceptibility classification of 11.0% of the modelling area experienced a variation; 6.3% of the area became less susceptible, while 4.7% experienced an increase in susceptibility.

The use of a generalized additive model was extremely useful to verify the model and in particular to attribute a physical interpretation to the fitted smooth functions. For the slope angle and the two climate variables, we had to constrain the behaviour of the fitted functions to avoid unexplainable local minima or maxima and trends.

The presented approach attaches a time reference to the susceptibility map (1991–2020) and demonstrates the utility of introducing process-related, non-stationary, physically plausible climate variables within a multi-temporal shallow landslide susceptibility analysis. Not only do the variables improve the model performance, but they make it adaptable to map the future evolution of landslide susceptibility including climate changes. The study represents the first step to achieve this aim and, besides the calculation of EAT_{ean} and ME_n for future time periods, it highlighted other possible developments for a precise and complete investigation of the impacts of climate change: firstly, the necessity to include higher spatial resolution rainfall data and secondly the possibility of introducing seasonal rather than all-year models.

7 Acknowledgements

We would like to thank the technical offices of Centro Funzionale Val d'Aosta, Struttura Attività Geologiche Val d'Aosta, and ARPA Val d'Aosta for data sharing and data management support. We would also like to thank Prof. Alexander Brenning for insightful discussion.

8 References

Aleotti, P., 2004. A warning system for rainfall-induced shallow failures. Engineering Geology 73, 247–265. https://doi.org/10.1016/j.enggeo.2004.01.007

Bajni, G., Camera, C.A.S., Apuani, T., under review. Deciphering triggering and preparatory factors for alpine rockfalls: a case study in Aosta Valley. Landslides.

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

Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. Scientific Data 5, 180214. https://doi.org/10.1038/sdata.2018.214

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

Bordoni, M., Galanti, Y., Bartelletti, C., Persichillo, M.G., Barsanti, M., Giannecchini, R., Avanzi, G.D., Cevasco, A., Brandolini, P., Galve, J.P., Meisina, C., 2020a. The influence of the inventory on the determination of the rainfall-induced shallow landslides susceptibility using generalized additive models. CATENA 193, 104630. https://doi.org/10.1016/j.catena.2020.104630

Bordoni, M., Vivaldi, V., Lucchelli, L., Ciabatta, L., Brocca, L., Galve, J.P., Meisina, C., 2020b. Development of a data-driven model for spatial and temporal shallow landslide probability of occurrence at catchment scale. Landslides. https://doi.org/10.1007/s10346-020-01592-3

Brabb, E.E., 1984. Innovative approaches to landslide hazard and risk mapping, in: Proc. 4th Int. Symp. Landslides. Canadian Geotechnical Society, Toronto, Canada, pp. 307–324.

Brenning, A., Schwinn, M., Ruiz-Páez, A.P., Muenchow, J., 2015. Landslide susceptibility near highways is increased by 1 order of magnitude in the Andes of southern Ecuador, Loja province. Natural Hazards and Earth System Sciences 15, 45–57. https://doi.org/10.5194/nhess-15-45-2015

Broeckx, J., Vanmaercke, M., Duchateau, R., Poesen, J., 2018. A data-based landslide susceptibility map of Africa. Earth-Science Reviews 185, 102–121. https://doi.org/10.1016/j.earscirev.2018.05.002

Büttner, G., Kosztra, B., Soukup, T., Sousa, A., Langanke, T., 2017. CLC2018 Technical Guidelines. European Environment Agency, Vienna, Austria.

Camera, C., Apuani, T., Masetti, M., 2015. Modeling the stability of terraced slopes: an approach from Valtellina (Northern Italy). Environ. Earth Sci. 74, 855–868. https://doi.org/10.1007/s12665-015-4089-0

Camera, C., Bruggeman, A., Hadjinicolaou, P., Michaelides, S., Lange, M.A., 2017a. Evaluation of a spatial rainfall generator for generating high resolution precipitation projections over orographically complex terrain. Stoch. Environ. Res. Risk Assess. 31, 757–773. https://doi.org/10.1007/s00477-016-1239-1

Camera, C., Zomeni, Z., Noller, J.S., Zissimos, A.M., Christoforou, I.C., Bruggeman, A., 2017b. 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

Carrara, A., Pike, R.J., 2008. GIS technology and models for assessing landslide hazard and risk. Geomorphology 94, 257–260. https://doi.org/10.1016/j.geomorph.2006.07.042

Chen, W., Li, Y., 2020. GIS-based evaluation of landslide susceptibility using hybrid computational intelligence models. CATENA 195, 104777. https://doi.org/10.1016/j.catena.2020.104777

Chen, W., Peng, J., Hong, H., Shahabi, H., Pradhan, B., Liu, J., Zhu, A.-X., Pei, X., Duan, Z., 2018. Landslide susceptibility modelling using GIS-based machine learning techniques for Chongren County, Jiangxi Province, China. Science of The Total Environment 626, 1121–1135. https://doi.org/10.1016/j.scitotenv.2018.01.124

Chung, C.-J.F., Fabbri, A.G., 2003. Validation of Spatial Prediction Models for Landslide Hazard Mapping. Natural Hazards 30, 451–472. https://doi.org/10.1023/B:NHAZ.0000007172.62651.2b

Citrini, A., Camera, C., Beretta, G.P., 2020. Nossana Spring (Northern Italy) under Climate Change: Projections of Future Discharge Rates and Water Availability. Water 12, 387. https://doi.org/10.3390/w12020387

Corno, I., 2019. Analisi di suscettibilità di frane superficiali in Valle d'Aosta attraverso il metodo statistico regressione logistica (Unpublished Master Thesis). Università degli Studi di Milano.

Dal Piaz, G.V., Bistacchi, A., Massironi, M., 2003. Geological outline of the Alps. Episodes Journal of International Geoscience 26, 175–180. https://doi.org/10.18814/epiiugs/2003/v26i3/004

Dal Piaz, G.V., Pennacchioni, G., Tartarotti, P., Carraro, F., Gianotti, F., Monopoli, B., Schiavo, A., 2010. Carta Geologica d'Italia alla scala 1: 50.000, Foglio 091 Chatillon.

Di Napoli, M., Di Martire, D., Bausilio, G., Calcaterra, D., Confuorto, P., Firpo, M., Pepe, G., Cevasco, A., 2021. Rainfall-Induced Shallow Landslide Detachment, Transit and Runout Susceptibility Mapping by Integrating Machine Learning Techniques and GIS-Based Approaches. Water 13, 488. https://doi.org/10.3390/w13040488

Dikshit, A., Sarkar, R., Pradhan, B., Jena, R., Drukpa, D., Alamri, A.M., 2020. Temporal Probability Assessment and Its Use in Landslide Susceptibility Mapping for Eastern Bhutan. Water 12, 267. https://doi.org/10.3390/w12010267

Dou, J., Yunus, A.P., Tien Bui, D., Merghadi, A., Sahana, M., Zhu, Z., Chen, C.-W., Khosravi, K., Yang, Y., Pham, B.T., 2019. Assessment of advanced random forest and decision tree algorithms for modeling rainfall-induced landslide susceptibility in the Izu-Oshima Volcanic Island, Japan. Science of The Total Environment 662, 332–346. https://doi.org/10.1016/j.scitotenv.2019.01.221

Ellero, A., Loprieno, A., 2018. Nappe stack of Piemonte–Ligurian units south of Aosta Valley: New evidence from Urtier Valley (Western Alps). Geological Journal 53, 1665–1684. https://doi.org/10.1002/gj.2984

Fang, Z., Wang, Y., Peng, L., Hong, H., 2020. Integration of convolutional neural network and conventional machine learning classifiers for landslide susceptibility mapping. Computers & Geosciences 139, 104470. https://doi.org/10.1016/j.cageo.2020.104470

Filippa, G., Cremonese, E., Galvagno, M., Isabellon, M., Bayle, A., Choler, P., Carlson, B.Z., Gabellani, S., Morra di Cella, U., Migliavacca, M., 2019. Climatic Drivers of Greening Trends in the Alps. Remote Sensing 11, 2527. https://doi.org/10.3390/rs11212527

Gariano, S.L., Guzzetti, F., 2016. Landslides in a changing climate. Earth-Science Reviews 162, 227–252. https://doi.org/10.1016/j.earscirev.2016.08.011

Gassner, Ch., Promper, C., Begueria, S., Glade, T., 2015. Climate change impact for spatial landslide susceptibility, in: Lollino, G., Manconi, A., Clague, J.J., Shan, W., Chiarle, M. (Eds.), Engineering Geology for Society and Territory. Springer International Publishing, Switzerland, pp. 429–433.

Geitner, C., Mayr, A., Rutzinger, M., Löbmann, M.T., Tonin, R., Zerbe, S., Wellstein, C., Markart, G., Kohl, B., 2021. Shallow erosion on grassland slopes in the European Alps – Geomorphological classification, spatio-temporal analysis, and understanding snow and vegetation impacts. Geomorphology 373, 107446. https://doi.org/10.1016/j.geomorph.2020.107446

Giordan, D., Cignetti, M., Wrzesniak, A., Allasia, P., Bertolo, D., 2018. Operative Monographies: Development of a New Tool for the Effective Management of Landslide Risks. Geosciences 8, 485. https://doi.org/10.3390/geosciences8120485

Glade, T., 2003. Landslide occurrence as a response to land use change: a review of evidence from New Zealand. CATENA 51, 297–314. https://doi.org/10.1016/S0341-8162(02)00170-4

Gobiet, A., Kotlarski, S., Beniston, M., Heinrich, G., Rajczak, J., Stoffel, M., 2014. 21st century climate change in the European Alps—A review. Science of The Total Environment 493, 1138–1151. https://doi.org/10.1016/j.scitotenv.2013.07.050

Goetz, J.N., Brenning, A., Petschko, H., Leopold, P., 2015. Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. Computers & Geosciences 81, 1–11. https://doi.org/10.1016/j.cageo.2015.04.007

Guyennon, N., Valt, M., Salerno, F., Petrangeli, A.B., Romano, E., 2019. Estimating the snow water equivalent from snow depth measurements in the Italian Alps. Cold Regions Science and Technology 167, 102859. https://doi.org/10.1016/j.coldregions.2019.102859

Guzzetti, F., Reichenbach, P., Ardizzone, F., Cardinali, M., Galli, M., 2006. Estimating the quality of landslide susceptibility models. Geomorphology 81, 166–184. https://doi.org/10.1016/j.geomorph.2006.04.007

Guzzetti, F., Reichenbach, P., Cardinali, M., Galli, M., Ardizzone, F., 2005. Probabilistic landslide hazard assessment at the basin scale. Geomorphology 72, 272–299. https://doi.org/10.1016/j.geomorph.2005.06.002

Hastie, T.J., Tibshirani, R.J., 1990. Generalized Additive Models. Chapman and Hall/CRC.

Herath, S., Wang, Y., 2009. Case Studies and National Experiences, in: Sassa, K., Canuti, P. (Eds.), Landslides – Disaster Risk Reduction. Springer, Berlin, Heidelberg, pp. 475–497. https://doi.org/10.1007/978-3-540-69970-5_25

Hungr, O., Leroueil, S., Picarelli, L., 2014. The Varnes classification of landslide types, an update. Landslides 11, 167–194. https://doi.org/10.1007/s10346-013-0436-y

Javadinejad, S., Dara, R., Jafary, F., 2020. Climate Change Scenarios and Effects on Snow-Melt Runoff. Civil Engineering Journal 6, 1715–1725. https://doi.org/10.28991/cej-2020-03091577

Kim, H.G., Lee, D.K., Park, C., Kil, S., Son, Y., Park, J.H., 2015. Evaluating landslide hazards using RCP 4.5 and 8.5 scenarios. Environ. Earth Sci. 73, 1385–1400. https://doi.org/10.1007/s12665-014-3775-7

Kjekstad, O., Highland, L., 2009. Economic and Social Impacts of Landslides, in: Sassa, K., Canuti, P. (Eds.), Landslides – Disaster Risk Reduction. Springer, Berlin, Heidelberg, pp. 573–587. https://doi.org/10.1007/978-3-540-69970-5_30

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, 217. https://doi.org/10.3390/geosciences10060217

Kotlarski, S., Keuler, K., Christensen, O.B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D., van Meijgaard, E., Nikulin, G., Schär, C., Teichmann, C., Vautard, R., Warrach-Sagi, K., Wulfmeyer, V., 2014. Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble. Geoscientific Model Development 7, 1297–1333. https://doi.org/10.5194/gmd-7-1297-2014

Lazoglou, G., Zittis, G., Anagnostopoulou, C., Hadjinicolaou, P., Lelieveld, J., 2020. Bias Correction of RCM Precipitation by TIN-Copula Method: A Case Study for Historical and Future Simulations in Cyprus. Climate 8, 85. https://doi.org/10.3390/cli8070085

Lombardo, L., Tanyas, H., 2020. Chrono-validation of near-real-time landslide susceptibility models via plug-in statistical simulations. Engineering Geology 278, 105818. https://doi.org/10.1016/j.enggeo.2020.105818

Lucas, D., Herzog, R., Iten, M., Buschor, H., Kieper, A., Askarinejad, A., Springman, S.M., 2020. Modelling of landslides in a scree slope induced by groundwater and rainfall. Int. J. Phys. Model. Geotech. 20, 177–197. https://doi.org/10.1680/jphmg.18.00106

Luino, F., De Graff, J., Roccati, A., Biddoccu, M., Cirio, C.G., Faccini, F., Turconi, L., 2020. Eighty Years of Data Collected for the Determination of Rainfall Threshold Triggering Shallow Landslides and Mud-Debris Flows in the Alps. Water 12, 133. https://doi.org/10.3390/w12010133

Maraun, D., Widmann, M. (Eds.), 2018. Reference Observations, in: Statistical Downscaling and Bias Correction for Climate Research. Cambridge University Press, Cambridge, pp. 87–95. https://doi.org/10.1017/9781107588783.008

Matiu, M., Petitta, M., Notarnicola, C., Zebisch, M., 2020. Evaluating Snow in EURO-CORDEX Regional Climate Models with Observations for the European Alps: Biases and Their Relationship to Orography, Temperature, and Precipitation Mismatches. Atmosphere 11, 46. https://doi.org/10.3390/atmos11010046

Menne, M.J., Durre, I., Vose, R.S., Gleason, B.E., Houston, T.G., 2012. An Overview of the Global Historical Climatology Network-Daily Database. Journal of Atmospheric and Oceanic Technology 29, 897–910. https://doi.org/10.1175/JTECH-D-11-00103.1

Mertens, K., Jacobs, L., Maes, J., Kabaseke, C., Maertens, M., Poesen, J., Kervyn, M., Vranken, L., 2016. The direct impact of landslides on household income in tropical regions: A case study from the Rwenzori Mountains in Uganda. Science of The Total Environment 550, 1032–1043. https://doi.org/10.1016/j.scitotenv.2016.01.171

Minasny, B., McBratney, A.B., 2016. Digital soil mapping: A brief history and some lessons. Geoderma 264, 301–311. https://doi.org/10.1016/j.geoderma.2015.07.017

Nahayo, L., Mupenzi, C., Habiyaremye, G., Kalisa, E., Udahogora, M., Nzabarinda, V., Li, L., 2019. Landslides Hazard Mapping in Rwanda Using Bivariate Statistical Index Method. Environmental Engineering Science 36, 892–902. https://doi.org/10.1089/ees.2018.0493

Nhu, V.-H., Shirzadi, A., Shahabi, H., Chen, W., Clague, J.J., Geertsema, M., Jaafari, A., Avand, M., Miraki, S., Talebpour Asl, D., Pham, B.T., Ahmad, B.B., Lee, S., 2020. Shallow Landslide Susceptibility Mapping by Random Forest Base Classifier and Its Ensembles in a Semi-Arid Region of Iran. Forests 11, 421. https://doi.org/10.3390/f11040421

Panziera, L., Gabella, M., Germann, U., Martius, O., 2018. A 12-year radar-based climatology of daily and sub-daily extreme precipitation over the Swiss Alps. International Journal of Climatology 38, 3749–3769. https://doi.org/10.1002/joc.5528

Peruccacci, S., Brunetti, M.T., Gariano, S.L., Melillo, M., Rossi, M., Guzzetti, F., 2017. Rainfall thresholds for possible landslide occurrence in Italy. Geomorphology 290, 39–57. https://doi.org/10.1016/j.geomorph.2017.03.031

Petley, D., 2012. Global patterns of loss of life from landslides. Geology 40, 927–930. https://doi.org/10.1130/G33217.1

Petschko, H., Brenning, A., Bell, R., Goetz, J., Glade, T., 2014. Assessing the quality of landslide susceptibility maps – case study Lower Austria. Natural Hazards and Earth System Sciences 14, 95–118. https://doi.org/10.5194/nhess-14-95-2014

Pistocchi, A., 2016. Simple estimation of snow density in an Alpine region. Journal of Hydrology: Regional Studies 6, 82–89. https://doi.org/10.1016/j.ejrh.2016.03.004

Ponti, S., Cannone, N., Guglielmin, M., 2021. A new simple topo-climatic model to predict surface displacement in paraglacial and periglacial mountains of the European Alps: The importance of ground heating index and floristic components as ecological indicators. Ecological Indicators 120, 106889. https://doi.org/10.1016/j.ecolind.2020.106889

Ponziani, M., Pogliotti, P., Stevenin, H., Ratto, S.M., 2020. Debris-flow Indicator for an early warning system in the Aosta valley region. Nat Hazards 104, 1819–1839. https://doi.org/10.1007/s11069-020-04249-5

Raffa, M., 2020. Il ruolo della neve nell'innesco di frane superficiali valutato con random forest: il caso del Mont Cervin e del Mont Emilius in Valle D'Aosta (Unpublished Master Thesis). Università degli Studi di Milano.

Ratto, S., Bonetto, F., Comoglio, C., 2003. The October 2000 flooding in Valle d'Aosta (Italy): Event description and land planning measures for the risk mitigation. International Journal of River Basin Management 1, 105–116. https://doi.org/10.1080/15715124.2003.9635197

Reichenbach, P., Busca, C., Mondini, A.C., Rossi, M., 2014. The Influence of Land Use Change on Landslide Susceptibility Zonation: The Briga Catchment Test Site (Messina, Italy). Environmental Management 54, 1372–1384. https://doi.org/10.1007/s00267-014-0357-0

Reichenbach, P., Rossi, M., Malamud, B.D., Mihir, M., Guzzetti, F., 2018. A review of statistically-based landslide susceptibility models. Earth-Science Reviews 180, 60–91. https://doi.org/10.1016/j.earscirev.2018.03.001

Samia, J., Temme, A., Bregt, A.K., Wallinga, J., Stuiver, J., Guzzetti, F., Ardizzone, F., Rossi, M., 2018. Implementing landslide path dependency in landslide susceptibility modelling. Landslides 15, 2129–2144. https://doi.org/10.1007/s10346-018-1024-y

Schilirò, L., Cepeda, J., Devoli, G., Piciullo, L., 2021. Regional Analyses of Rainfall-Induced Landslide Initiation in Upper Gudbrandsdalen (South-Eastern Norway) Using TRIGRS Model. Geosciences 11, 35. https://doi.org/10.3390/geosciences11010035

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., Piciullo, L., Gariano, S.L., 2018a. A review of the recent literature on rainfall thresholds for landslide occurrence. Landslides 15, 1483–1501. https://doi.org/10.1007/s10346-018-0966-4

Segoni, S., Tofani, V., Rosi, A., Catani, F., Casagli, N., 2018b. Combination of Rainfall Thresholds and Susceptibility Maps for Dynamic Landslide Hazard Assessment at Regional Scale. Front. Earth Sci. 6. https://doi.org/10.3389/feart.2018.00085

Shou, K.-J., Lin, J.-F., 2020. Evaluation of the extreme rainfall predictions and their impact on landslide susceptibility in a sub-catchment scale. Engineering Geology 265, 105434. https://doi.org/10.1016/j.enggeo.2019.105434

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

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

Stevenazzi, S., Masetti, M., Nghiem, S.V., Sorichetta, A., 2015. Groundwater vulnerability maps derived from a time-dependent method using satellite scatterometer data. Hydrogeol J 23, 631–647. https://doi.org/10.1007/s10040-015-1236-3

Stumvoll, M.J., Canli, E., Engels, A., Thiebes, B., Groiss, B., Glade, T., Schweigl, J., Bertagnoli, M., 2020. The "Salcher" landslide observatory—experimental long-term monitoring in the Flysch

Zone of Lower Austria. Bull. Eng. Geol. Environ. 79, 1831–1848. https://doi.org/10.1007/s10064-019-01632-w

Subramanian, S.S., Fan, X., Yunus, A.P., Asch, T. van, Scaringi, G., Xu, Q., Dai, L., Ishikawa, T., Huang, R., 2020. A Sequentially Coupled Catchment-Scale Numerical Model for Snowmelt-Induced Soil Slope Instabilities. Journal of Geophysical Research: Earth Surface 125, e2019JF005468. https://doi.org/10.1029/2019JF005468

Szumilas, M., 2010. Explaining Odds Ratios. J. Can. Acad. Child Adolesc. Psychiatry 19, 227–229.

UNISDR, 2015. Making development sustainable: the future of disaster risk management, Global Assessment Report on Disaster Risk Reduction. UNISDR, Geneva, Switzerland.

van Westen, C.J., Castellanos, E., Kuriakose, S.L., 2008. Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. Engineering Geology 102, 112–131. https://doi.org/10.1016/j.enggeo.2008.03.010

Wood, S., 2020. Package “mgcv” version 1.8-33 - Mixed GAM Computation Vehicle with Automatic Smoothness Estimation.

World Meteorological Organization (WMO), 2007. The role of climatological normals in a changing climate, WCDMP No. 61, WMO/TD No. 1377. WMO, Geneva.