Cignetti *et al. Geoenvironmental Disasters* (2024) 11:36
https://doi.org/10.1186/s40677-024-00300-w

Geoenvironmental Disasters

RESEARCH

Open Access

A-posteriori analysis of the performance of a rockfall susceptibility map

M. Cignetti[1†], D. Godone[1*†], D. Cardone[1,3], D. Giordan[1] and D. Bertolo[2]

Abstract

Background Rockfalls pose a serious threat along the main road network, representing a major hazard in mountainous territory and causing damage and victims. Currently, susceptibility mapping represents a starting point to identify areas more susceptible to rockfall occurrence, a key approach in land use planning and risk management. Despite the extensive use of these maps by decision makers and administrators, the usability of such maps over time and their reliability represent a poorly discussed and examined feature.

Methods Here, we proposed a-posteriori analysis of a three-year-old rockfall susceptibility map, generated along the main road network of the Aosta Valley, an alpine region of north-western Italy. To verify map consistency over time, we implemented a dual-analysis in GIS-environment and by text mining, to respectively analyse the geocoded data and textual information derivable from the regional landslide inventory. The first one allowed us to extract rockfall events occurred after the susceptibility map generation. By this way, we operated to spatially and temporally verify the map consistency. Jointly, the textual information reported in the Event Description Form, linked to each geocoded event, are being exploited. This allowed us to derive relevant information about occurred damage and their degree, presence of protective measures or secondary roads, i.e. involvement of, farm or forestry, road.

Results The implemented approach allowed us to prove the quality of the previous map in terms of reliability, robustness and degree of fitting respect to the succession of rockfall occurrence over time. After only three years as many as 198 rockfall events have been occurred and collected since the map was generated. Particularly, 80% of rockfall fit with "high" and "very high" susceptibility classes, and pointed out large involvement of main roads in rockfall occurrence, representing the most affected target with damage to road pavement and vehicles, as well as a relevant involvement of existing rockfall barriers and of the dense network of forestry roads and footpaths that characterize this alpine region.

Results The proposed approach representing a starting point in landslide susceptibility map verification and usability as valid instrument for a reliable land planning.

Keywords Regional landslide inventory, Aosta valley region, GIS-environment, Map reliability assessment, Mountain area, Road network, Text mining

†M.Cignetti and D. Godone contributed equally to this work.

*Correspondence:
D. Godone
danilo.godone@irpi.cnr.it

[1]National Research Council of Italy, Research Institute for Geo-Hydrological Protection (CNR IRPI), 10135 Torino, Italy
[2]Strutture Attività Geologiche, Regione Autonoma Valle d'Aosta, 11020 Quart, Italy
[3]Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, Netherlands

Introduction

Rockfall events are extremely widespread in mountain territories worldwide, representing a major hazard that causes considerable damage and loss of life. Due to their high energy, long runout distance and rapidity, these unpredictable phenomena represent a relevant threat, mainly to road and railway networks (Guzzetti et al. 2004; Palma et al. 2012), requiring specific risk reduction operative strategies for a proper management of the associated risk and increase safety of the people against instability processes (Giordan et al. 2020; Clemente et al. 2023). Rockfall hazard estimation is a relevant topic in literature, currently representing a challenge to establish the associated risks. The frequency assessment of annual rockfall events, of a given magnitude, is preparatory to a detailed hazard computation (Hungr et al. 1999). Without a strict definition of the magnitude-frequency relationship, the rockfall hazard estimation is not feasible. In those cases, susceptibility estimation may represent a rapid way to preliminarily identify the potential critical sites prone to rockfall events. The susceptibility corresponds to the propensity of an area to generate landslides on the basis of geo-environmental conditions and historical records (Brabb 1984; Guzzetti et al. 2005). Considering the assumption that landslides will occur under the same conditions, which led to past and present instability (Varnes 1984; Carrara et al. 1991; Furlani and Ninfo 2015), susceptibility definition may ensure to identify areas where landslides are expected. Consequently, the territory zonation may represent an important task for the prioritization and plan structural protective and mitigation measures, suitable for regional and local authorities in risk management. In the scientific literature, numerous approaches and methods have been proposed for rockfall susceptibility assessment, based on diverse qualitative and quantitative approaches, heuristic (Powell et al. 2002; Raetzo et al. 2002), statistic (Shirzadi et al. 2012) and trajectory-energy/velocity approach (Guzzetti et al. 2002; Alvioli et al. 2021). Qualitative approaches require less input data, derived from field information or by overlaying indexed maps layers, often with the attribution of a weight, but suffer the subjectivity of the operator. Conversely, quantitative methodologies are based on numerical values related to the occurrence of the analysed phenomena in the area of interest (Guzzetti et al. 1999), representing the preferred approach to be applied. Considering large areas, an extensive data collection can represent a challenge in terms of quality and quantity of available information to perform a consistent quantitative analysis. Therefore, a qualitative susceptibility assessment might be preferable for regional investigation, where an accurate assessment of frequency of rockfall events, as well as data acquisitions about travel distance and volume of the detached mass, are often incomplete and discontinuous.

Susceptibility maps generally represent an effective instrument to support public administrations, regional and local authorities and decision-makers, to identify those sectors with structures and infrastructures at risk, operating for a proper urban and infrastructures planning. Nowadays, susceptibility maps are a well-known tool in land use planning, in landslide-prone territories, and despite a proper validation procedure of the utilized model, and its results, at the end of its computation, the efforts to critically evaluate their usability and reliability over time are limited. Usually, the evaluation of the susceptibility models includes a validation of the map, the model fitting and prediction performances. Regarding the validation, in literature, diverse validation analysis are applied to assess the performance of a susceptibility map, by applying temporal, spatial or random approaches (Reichenbach et al. 2018). In spatial and random strategies, the validation is based respectively on a subset of territory and a random geographic selection of territory. Instead, in temporal approaches, two reference sets of landslide inventories, typically occurred before and after a certain date, are considered for the validation, and in any case with reference to a limited time interval. Considering the prolonged use of a susceptibility map over time, with both spatial or temporal validation approaches the representativeness of the susceptibility map can suffer of time-dependence. Also regarding the fitting and prediction performance, typically verified by landslide density or frequency, success rate curve or Receiver Operating Characteristic (ROC) curve (Reichenbach et al. 2018), the validation of a susceptibility map and its consistency is obtained over the period considered, while its reliability over time remains a relevant issue currently poorly debated in the scientific field.

In this work, we developed an approach to verify the representativeness of a three-year-old rockfall susceptibility map, developed by the authors (Cignetti et al. 2020) along the road network of the Aosta Valley Region (North-western Italy). Leveraging on this existing rockfall susceptibility map, validated by analysing the rockfall events inventoried in the regional catalogues, and showing a good prediction performance verified by ROC curve, an analysis of the consistency of the model after several years was implemented. Taking advantage of the available regional landslide inventory, i.e. the '*Catasto Dissesti*' (CD) inventory, dual-format data were derived. The geocoded information, in shapefile format, was managed in a GIS-environment, in order to collect and organize the rockfall phenomena occurred in the regional territory, preparatory step for the evaluation of the map-reliability over time. Jointly, the large number of textual information derivable from the 'Event Description Form'

(EDF), a document predisposed by the regional administrators for the description of each occurred phenomenon, available in pdf format, was analysed through text mining approach, to extract information about eventual damage occurrence, presence of protective measures or secondary, i.e. farm or forestry, road involvement. Through this dual-analysis, we operated to establish the quality in terms of reliability, robustness and degree of fitting of the susceptibility map with the succession of the occurrence of such events over time. The prediction capability evaluation of a susceptibility map to identify the propensity of slope to fail is not trivial, opening a useful discussion about the correct usage of these maps over time, an essential feature for a proper technical protection measure scheduling and a suitable land use planning.

The Aosta Valley Region, Western Italian Alps

The Aosta Valley Region (AVR) is a small alpine region (about 3200 km^2) located in the Western Italian Alps (Fig. 1). Elevation in this mountainous territory ranges from 300 m to 4000 m above the sea level, with more than a half of the territory above 2000 m a.s.l., and peaks higher than 4000 m (e.g., Mont Blanc, 4810 m, Monte Rosa, 4635 m and Monte Cervino, 4478 m) (Fig. 1a). Compared to other mountainous regions worldwide, this alpine territory is densely urbanized and populated (in 2023 more than 123 thousand inhabitants, (ISTAT 2023), despite the high altitudes and the complex morphology. A dense road network extends to the entire main valley and branches into the lateral valleys, with a total length of 2600 km, including highway, state, regional and municipal roads. Moreover, an extensive and widespread network of farm, forestry roads and footpaths extend

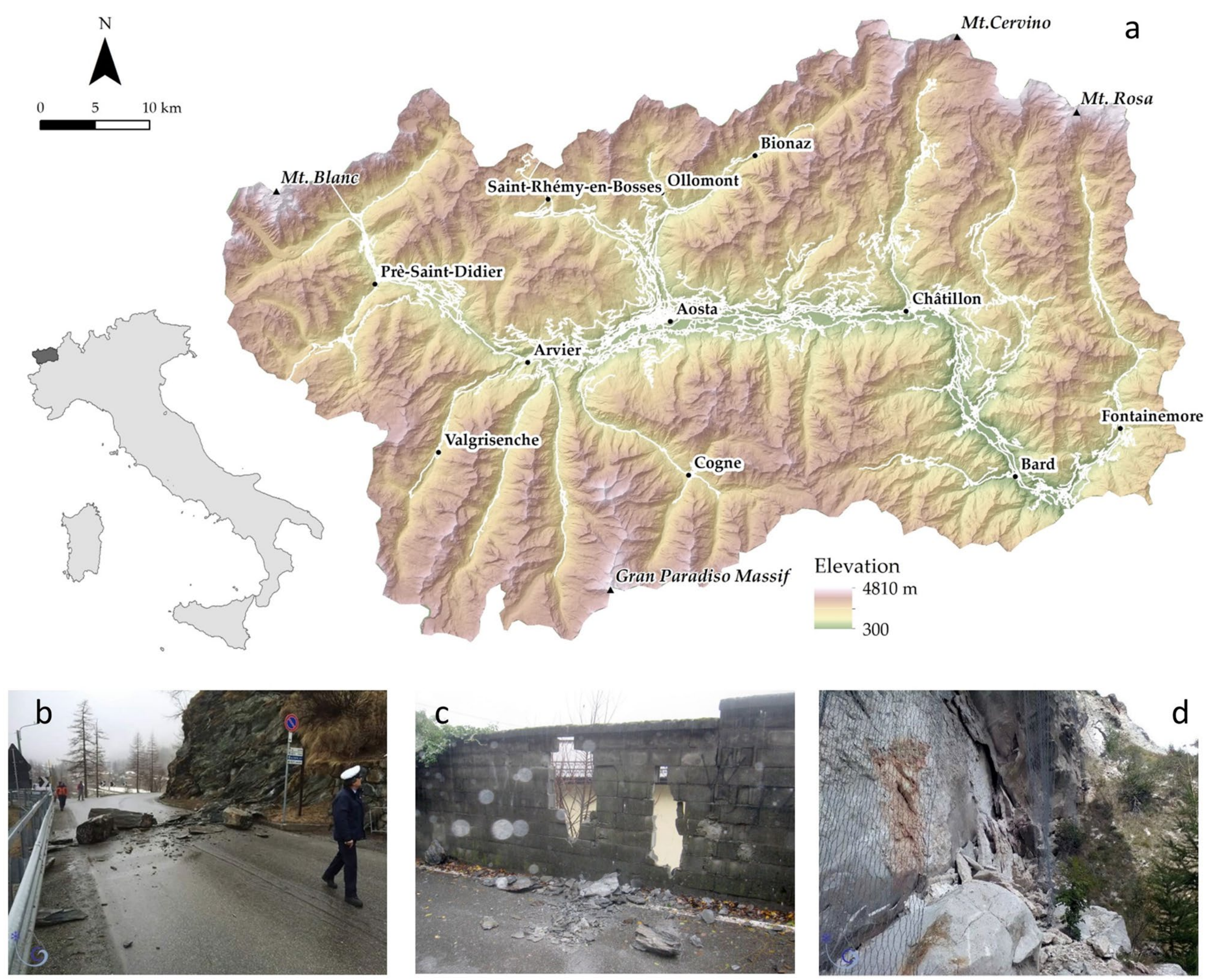

Fig. 1 (**a**) Relief terrain map of the Aosta Valley mountainous territory (north-western Italy), associated to the elevation map, showing the extended road network (white polylines), branched along the main valley of this alpine region highly affected by rockfalls that cause relevant damage to (**b**) road pavement and vehicles, (**c**) urban settlements, and (**d**) existing structural protective measures

throughout the entire territory, representing a frequently travelled route mainly for agricultural/forestry purposes and forest fire prevention, as well as for tourist purposes. The main urban areas are located at the valley bottom of the main one, while along the valley sides, small urban agglomerations, e.g. villages, hamlets, up to scattered buildings are more prevalent. Among these small villages and hamlets, many sites with great historical and artistic value, such as castles or fortresses are present.

The complex geological setting corresponds to the axial zone of the Western Alps, consisting of an imbricated stack of continental and oceanic metamorphic domains, passing through the European-vergent, i.e. Australpine Pennidic and associated Ophiolitic Units domains, and, in the northwestern marginal sector, the Ultrahelvetic and Helvetic units. This complex pile of nappes, characterized by a post-collisional tectonic activity, shows a neo-tectonic dislocation system activation, mainly represented by the Aosta-Ranzola semi-graben (Bistacchi et al. 2001). This compound structural-geological setting highly influenced the relief evolution and slope dynamic.

Among the main morphodynamic agents, glacial dynamic had played a central role, testified by the extended large main Dora Baltea valley, converging in the impressive Ivrea moraine amphitheater, remarkable evidence of the Quaternary glaciations (Gianotti et al. 2008). Glacial morphodynamic influenced the past and current dynamics, by modelling depositional and erosional landforms (Carraro and Giardino 2004). The geomorphic activity of watercourses superimposes the glacial forms, causing a progressive erosional deepening of the valleys, setting large alluvial and mixed fans widespread in the valley bottom, mainly created by debris flows (Ratto et al. 2003). Because of the high reliefs and the complex geological and structural setting of this mountain area, the AVR is a territory highly prone to landslides (Palomba et al. 2015). Gravitative phenomena widely affect the entire regional territory, vary in size and type, from large slow-moving landslides to rapid and impulsive debris flow, earthflow, and rockfall phenomena. According to the regional landslide inventory, i.e. IFFI project - *Inventario Fenomeni Franosi Italiano* (Trigila et al. 2008), rockfalls represent one of the most frequent landslide types, more than 29% (Giardino and Ratto 2007). The widespread distribution within the regional territory entails a high risk of rockfalls, frequently affecting the urban settlements and the extended main road network (Giordan et al. 2021), causing severe damage and victims. Consequently, these rapid and abrupt phenomena cause significant injuries to road pavement and vehicles (Fig. 1b), as well as to buildings (Fig. 1c) and structural protective measures (Fig. 1d), causing some victims along the communication routes branched along the main and secondary valleys.

Materials and methods

In this work, a study for the assessment of the performance and reliability of a rockfall susceptibility map over time has been carried out. Focusing on the AVR, an alpine territory extremely affected by these abrupt and frequent phenomena, a representative case is guaranteed. Taking advantage of the three-years-old rockfall susceptibility map generated along the main road network of this region (Cignetti et al. 2020), and exploiting the large amount of data derived from the long-established operational procedure applied by the regional authorities (Giordan et al. 2021), the map-reliability survey was implemented. Figure 2 shows the conceptual scheme of the overall procedure proposed, based on a mutual analysis of the dual-format data of the regional landslide inventory: (i) vector data of geo-coded rockfall events, analysed in a GIS-environment; (ii) textual information derived from the EDF predisposed by the regional authorities, investigated by a text mining approach in R environment.

Pre-existing rockfall susceptibility map

A rockfall susceptibility map existing for the AVR territory has been computed by the authors (Cignetti et al. 2020), considering a buffer of 250-m along the entire regional road network (Fig. 3).

This rockfall susceptibility map was obtained by the application of the Analytic Hierarchy Process (AHP) approach (Saaty 1980), a semi-quantitative index-based method founded on multi-criteria decision analysis. Through this approach, a multi-objective comparison allows to define a scale of preference from a set of alternatives. A pair-wise comparison matrix of the six causative factors considered, i.e., slope, aspect, elevation, lithology, land use and distance to road, as well as for each class of factor, was constructed. Thus, for both causative factors and to the classes defined for each factor, is assigned a weight. The judgments were assigned following subjective expert-based decisions on the relative importance of each factor. To verify the consistency of the judgments, the Consistency Ratio (CR) was computed. The results of the CR computed for each factor show values lower than 10%, thereby the consistency of the judgments is confirmed. This map covers an area of interest of about 475 km^2 along the road network, of which more than 40% displays very high susceptibility to rockfalls, excluding the flat-gently slope areas. This susceptibility map consistently matches the previous inventoried rockfall events, referring to the regional landslide inventory CD, showing a good predictivity respect to the available observation time from 1180 up to February 2020. On a total of 1654 rockfall phenomena, 70% of them occurred in correspondence of areas with high and very high susceptibility, characterized by steep slopes and medium elevations,

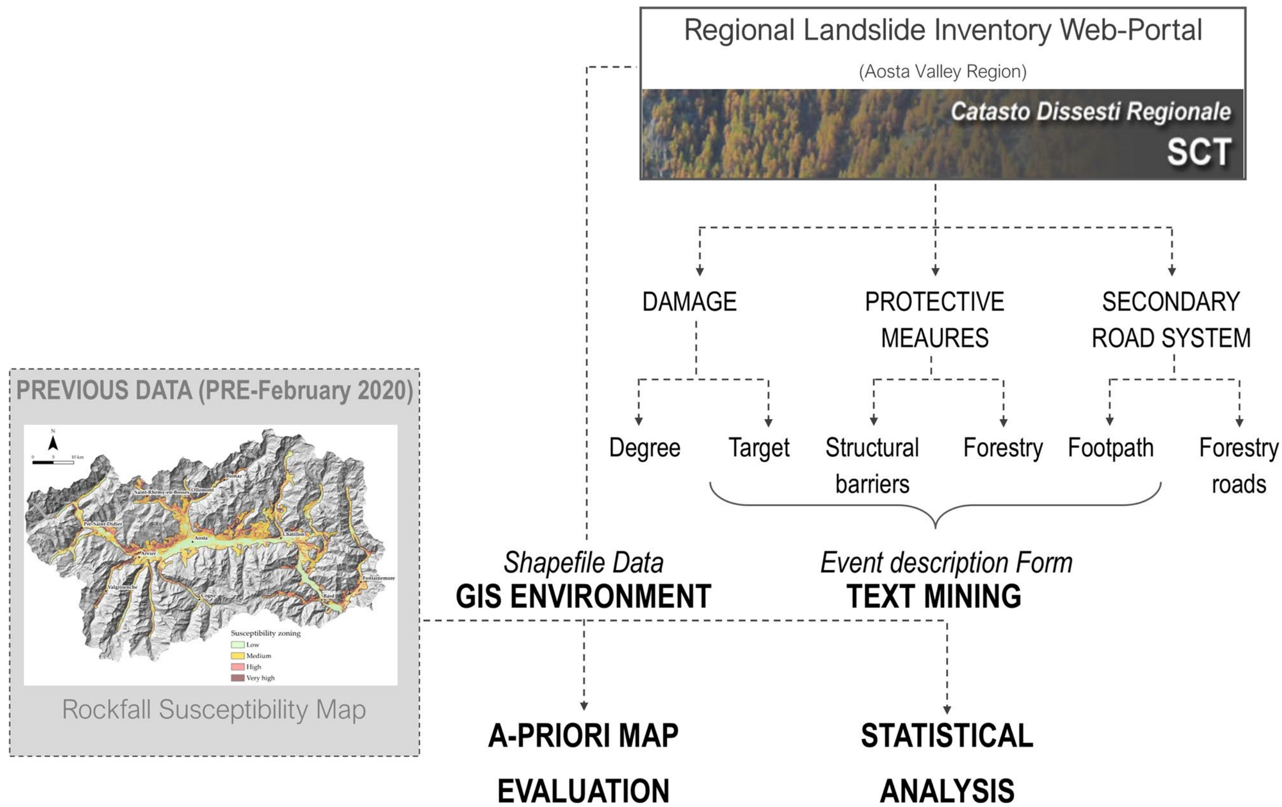

Fig. 2 Conceptual scheme of the map-reliability investigation

and mainly in correspondence of schistose lithologies of the Piemonte-Liguria Ocean crust and Mesozoic coverage sediment pertaining to the Piedmont Zone, in shrub and/or herbaceous vegetation areas (45%) or woodlands (36%). Further validation was carried out by ROC curve application, obtaining a success rate of 75%, confirming the reliability of the obtained map. Thus, the obtained map, hereinafter pre-February 2020 Rockfall Susceptibility Map, was validated and considered reliable for the time of its creation.

Regional landslide inventory web-portal

In Italy, an extensive and detailed classification of landslides was carried out with the compilation of an inventory called "*Inventario Fenomeni Franosi Italiani*", the IFFI Catalogue (Trigila et al. 2008). This Italian Landslides Inventory is made available from the administrators of each region, providing a regional classification of all the type of landslides based on standardized criteria able to provide a homogeneous geocoded database. For the area of interest, this inventory currently converges into the CD, i.e. an on-line regional catalogue managed by the regional authorities (Centro Funzionale Regione Autonoma Valle d'Aosta 2020). This on-line portal guarantees a valuable collection and storage of landslides occurred, through a constant updated operated by a dedicated IT procedure, implemented by the regional authorities, together with the Aosta Valley Forestry Corps and Fire Brigade. Every time a rockfall occurs, this operational procedure, applied at regional scale, is activated, consisting of the survey by the operators, which document the phenomenon filling the EDF predisposed in a standard version by the regional authorities. Promptly, the regional technicians revise the document, i.e. an EDF in pdf format for each occurred event, and constantly update the regional inventory. Within the forms compiled for each phenomenon, the following information is given: (i) code of the inventoried phenomenon, (ii) typology, (iii) date and time, (iv) a brief description, (v) triggering factor (if known), (vi) (estimated) size/morphometry, (vii) damage, (viii) geographical references. Moreover, on-field survey documented by photographs of the area involved are carried out, providing a useful overall analysis for risk management and prevention purposes. Currently, the rockfalls inventoried in the CD cover a quite extended time span, with the first recoded events dating back to 1180. The oldest phenomena derived mainly from indirect sources (e.g., newspaper, verbal witnesses), show a partially completed EDF. Accuracy increase progressively for the phenomena occurred in the last 40 years, overall ensuring a large amount of data and information with a standardized acquisition procedure.

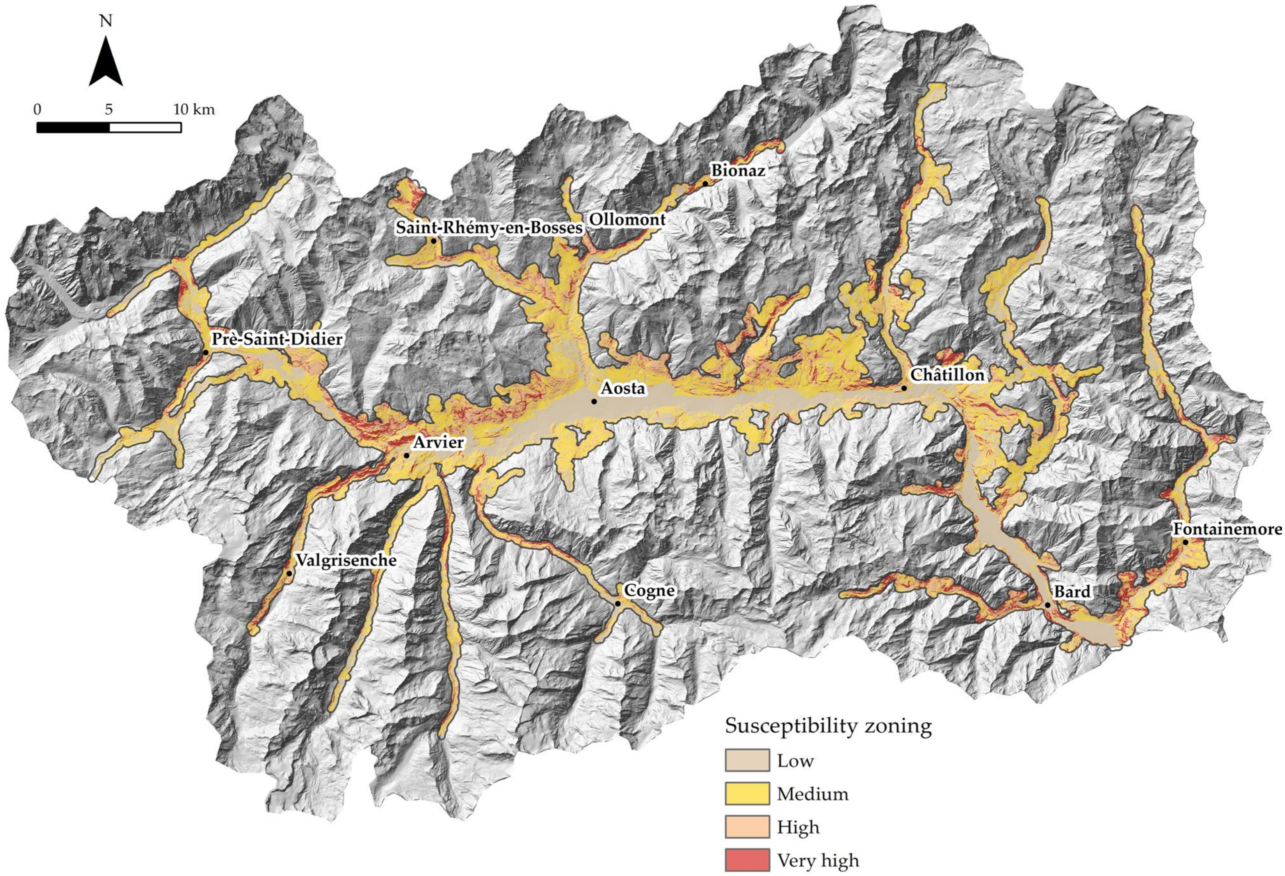

Fig. 3 Rockfall Susceptibility Map along the main road ways. (modified from Cignetti et al. 2020)

Susceptibility map reliability examination

By exploiting the data completeness derivable from the regional landslide inventory, two main type of datasets are provided: (i) vector data, in shapefile format, providing the geo-localization of the collected phenomena, (ii) 'EDF', in pdf format, reporting textual information to properly define the occurred phenomenon.

Vector data consists of geometric entities, i.e., points, line and polygons representing the rockfall phenomena depending to the size. Currently, 1712 rockfalls are represented as points, variably geo-localized in correspondence of the source area or of the block's trajectory end point, on the base of the survey, observation and accessibility of the area, and 619 as polygons, scattered throughout the regional territory. The shapefiles are made available by the regional authorities, and currently available under request. Each geometry reports a unique event-code, linked to the corresponding EDF, and few others code and information, e.g. type of movement and municipality. Operating in a GIS-environment, the geometries related to rockfall phenomena, occurred after the pre-February 2020 Rockfall Susceptibility Map realization, were selected over an observing period from February 2020 to December 2023. A preliminary analysis of the sub-selected rockfalls distribution is carried out to spatially verify the correspondence with susceptibility classes of the existing map. By exploiting the main DTM-derivative products from the 2 m-DTM (https://geoportale.regione.vda.it/download/dtm/), the geo-lithological information, mainly derived from CARG project (ISPRA Ambiente 2012), i.e., Geological and Geothematic Mapping, and land use cover (Regione Autonoma Valle d'Aosta 2020, https://geoportale.regione.vda.it/download/carta-copertura-suolo/), a further contextualization of the variables mainly affecting the occurrence of rockfall is carried out.

In addition to quantitative maps evaluations, a systematic evaluation of the EDFs, in Italian language, downloadable from the CD web-portal, was accomplished. The regional web portal, while a valuable source of information, is not a database in strict sense, and therefore cannot be interrogated by query in an automatic and unique way. The extraction of narrative information shall be used to evaluate the consistency of the map, focusing on some relevant information about rockfall occurrence and related effects. Through a text mining approach an automatically discover, retrieve and extraction of meaningful information in the corpus of text, was implemented.

Using the statistical open source environment R (Core Team 2022), a total of 1702 EDF, over 2331 event collected in the regional inventory, were considered and analyzed. The forms analyzed cover a time span of 23 years (from the 1st of January 2000 to 13st December 2023), corresponding to those events with an in-depth description and en extensive compilation of the EDF. In fact, after the extreme October 2000 flood event (Ratto et al. 2003) that widely affected the entire region, a growing need for a tool to characterize the territory for an effective risk assessment and management, led to a more complete and extensive data collection. Instead, for rockfalls occurred before 2000, usually a more concise compilation is observed, mainly derived from indirect information and with partial/incomplete filling of the fields in the form. Therefore, the EDFs with date before 2000 were not considered for the text mining analysis. By text mining, we were able to process unstructured and rough textual data, with the aggravating factor of the PDF format, by combining approaches involving linguistics, statistics and computer sciences (Salloum et al. 2018). The main purpose is to convert the textual and narrative information into a list of organized elements, for a structured and hierarchical representation of the data of each rockfall phenomena. The high-quality derived information are focused on three main categories:

i. damage;
ii. protective measures,
iii. secondary road system.

The information extraction (IE) primarily focuses on the identification of recurrent keywords, defined by an expert based approach, aimed at defining the existence of damage, the presence of both structural and natural protective measures (e.g., rockfall barriers, protection forest), and the potential involvement of the secondary road system in rockfall occurrence. A progressive hierarchization of the information was then realized to identify entities that co-occur within the text, i.e. to extract relationships of a certain type, as 'damage-degree of damage', 'damage-target', 'secondary road system-forestry-roads reached', 'protective measurement-degree of forestry protection'. The IE procedure was carried out to structure the main categories designed, considering the term concerning the specific category, as well as the potential term variation and synonyms, for example<impact>, < impacting>and <impacted>, can be considered as keywords to search presence of damage, or <tree>, < plant>and <spruce>to identify, in the event description, the involvement of a forest in rockfall occurrence. All the keywords considered to identify the three main categories and those to hierarchize derived information are summarized in Fig. 4. A detailed list of all original Italian term considered, associated to their translation, are provided in Supplementary Information.

Results and discussion

GIS-based analysis of the pre-february 2020 Rockfall susceptibility map

The pre-February 2020 Rockfall Susceptibility Map computed by Cignetti et al. 2020 along the roads of the AVR has shown the possibility to obtain, through a rapid and repeatable procedure, a valid tool for the zonation of the area more susceptible to rockfalls, at regional scale. However, an evaluation of the reliability and validity of a susceptibility map in time remain an open issue. Three years after the production of the pre-February 2020 Rockfall Susceptibility Map, validated and tested in an alpine region severely subject to rockfall events, an audit of its consistency over years represents an unexplored field. The availability of detailed data and information about events that occurred subsequently, collected in a standard mode and verified by regional authorities, represents a key factor to analyse over time the reliability of a susceptibility map. After February 2020, a total of 198 rockfall events in just three years have been recorded, of which 145 affecting the main roads of the AVR. To verify whether the pre-February 2020 Rockfall Susceptibility Map (Fig. 3) may be considered as satisfactory, a comparison of the post-February 2020 rockfall events with the AHP classes of the existing susceptibility map was carried out. This analysis shows that 76% of the rockfall collected after February 2020 fall in areas with very high (39%) and high susceptibility (37%), while less than 17% in area with medium susceptibility, and slightly more than 5% in areas with low susceptibility. Figure 5 shows the map of the distribution of the rockfall events occurred after February-2020 and the related spatial density. the highest density (8–9 and >9 events) is observed in the sector near to Bard village, at the beginning of Valtournenche, near Chatillon, and in the sector near to Avise town. Rockfall events with high and very high susceptibility predominantly occurred in areas with high slope gradient, while there appears to be no degree of dependence with elevation factor, mainly variables between 500 and 1500 m a.s.l., confirmed what stated in Cignetti et al. 2020. It is important to note that a degree of subjectivity in the description of events by regional operators may occur, as non-unique criteria are used for the compilation of the event description (e.g., inability to identify the source area resulting in the location of the event in correspondence of accumulation area or of impact sector). Analysing the lithology, most of the rockfall occurred in schistose rocks, mainly represented by calcschists, serpentinoschists, micaschists, while considering the land use aspect, the post-February 2020 events mainly occurred in areas occupied by shrub and/or herbaceous

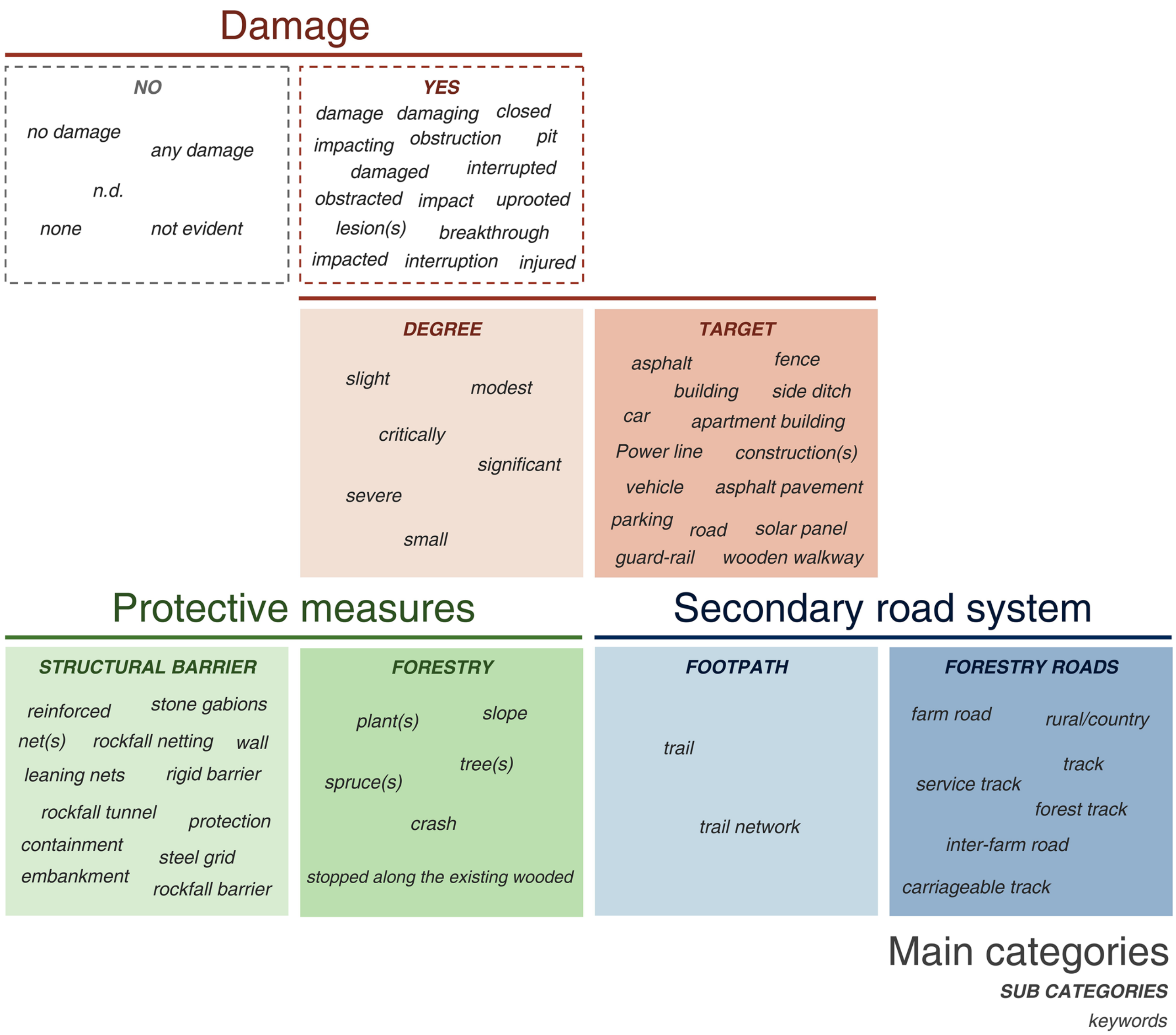

Fig. 4 Summary of the keywords, from the three main categories, selected for the text mining application

vegetation, or in correspondence of woodlands, confirming the results pointed out by the pre-February 2020 Rockfall Susceptibility Map (Cignetti et al. 2020).

Text mining outputs analysis

The implemented methodology, apart from the vector data consideration, jointly exploited the vast amount of data and information collected over more than 20 years in the regional catalogue CD, to assess the reliability of the existing susceptibility map. The search over 1702 EDF, considering a time span from 2000 to 2023, yielded the preliminary extraction of narrative information about the key noteworthy evidences on rockfall occurrence and related damage. This analysis used written narratives form in pdf format. Commonly, in literature, text mining is intended to automatically retrieve, and extract information in a corpus of text (Doan Duong et al. 2016; Yagoub et al. 2020). mainly csv file or html. The use of precompiled form only available in pdf format, posed a further challenge in spite of analyzing plain text sources as the file readability is more difficult to implement in an automatic procedure.

Our process of identification of hierarchical information, about the occurred rockfalls of the AVR, was based on a set of objectives for connecting archival data to high-quality information that, automatically, synthesizes and optimizes the data contained and its accessibility, in a way that maximizes the utility to characterize the rockfall events, their evolution and the associated risk. In large amount of data, such as in the case of the regional landslide inventory of the AVR, the continued growth of information over time suggest that a manual

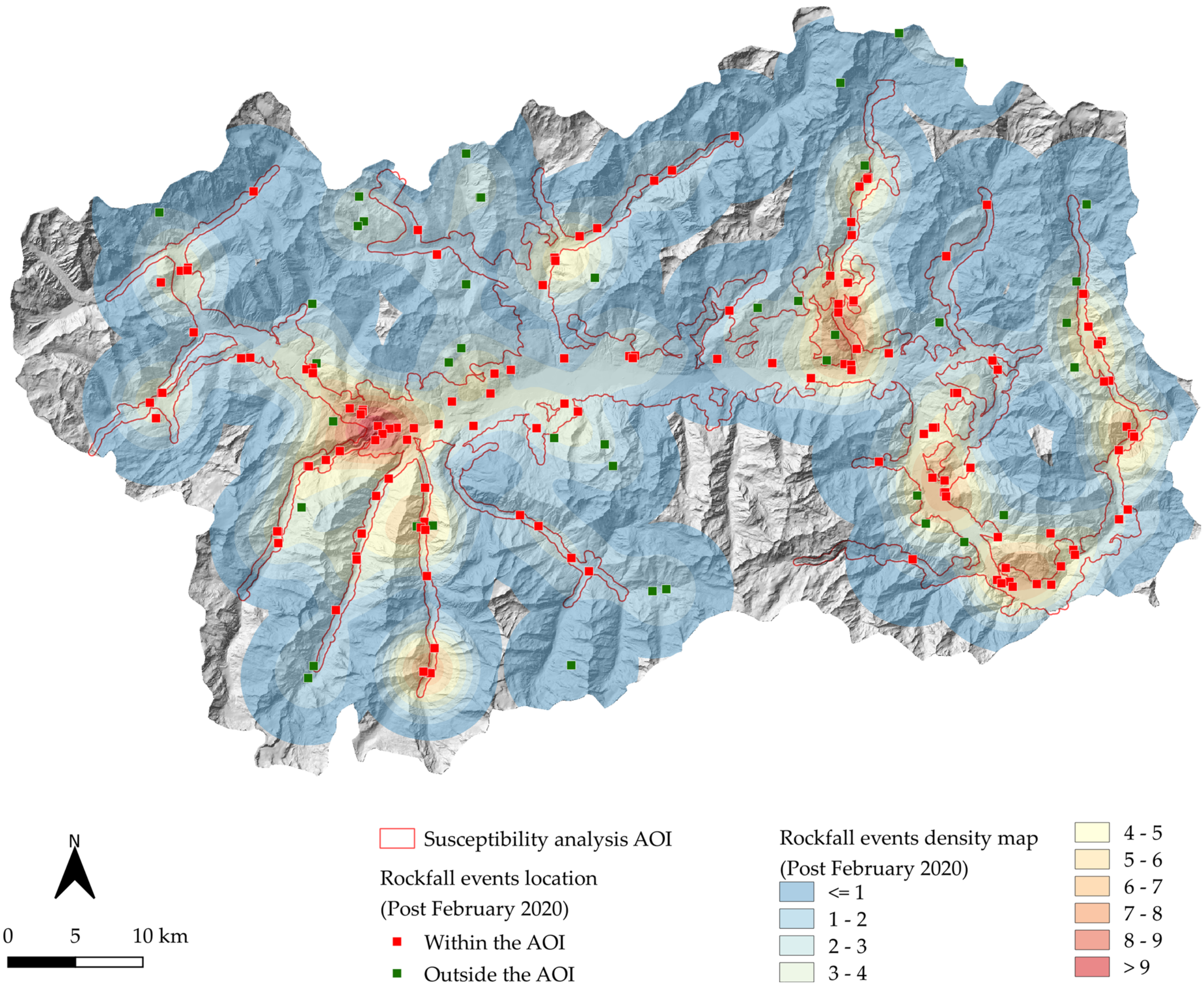

Fig. 5 Density map and location of the rockfall events occurred after February-2020 within the regional territory

examination of this type of archive can be incredibly difficult, time consuming and challenging to process task even for expert users. Though the automated application of text mining in R language, we are capable of providing effective interactions with valuable archival content to produce useful summarizations and visualizations about rockfall event, their features and impacts, affecting the area of interest. The main obtained results are summarized in the following sub-section, focusing on the three main categories: damage, protective measures and secondary road system.

Damage information extraction

Information about impacts extracted from the EDF, include the identification of the occurrence of damage, the related degree of damage, and of the affected target. The investigation concerns only the main anthropic elements, e.g., regional and national roads, buildings, power lines, as the most sensitive elements in term of risk and having the highest impact on population in case of partial or complete destruction. Results from the text mining allowed to quantify the frequency of occurrence of the, previously, selected keywords (Fig. 4), in terms of their appearances in the text. Primarily, we extracted the keywords assigned to the identification of the presence or absence of damage, recognizing that slightly less than 16% of the analysed forms reported terms related to the presence of damage. Those cases reporting an occurred damage or impact were further analysed to determine the degree of damage, distinguishing for three classes of damage: 42% of impacts with severe degree of damage, 14% moderate and 44% slight degree of damage. Another relevant relationship between damage and hit target was extracted, again on the basis of the previously selected keywords referring to the main anthropic structures and infrastructures. Specifically, six sub-classes, i.e., roads; vehicles; people; cattle, buildings; and facilities, are designed to adequately subdivide the keywords collected on account of the major target affected. Apart from the most common structures and infrastructures, such as road, houses, parking, power line and so on, we focused on those sub-groups of keywords that referred to the involvement of people and vehicles, as element at risk of primary interest. Moreover, we considered the involvement of farms animals, given the economic relevance and the widespread of agriculture and pastoralism activities in a mountainous territory such as the AVR. Figure 6 summarizes the frequencies related to each identified target. The obtained results showed that, among the recognized targets, the most frequent corresponds to the road sub-class targets (37.1%), primarily represented by road network target and then by other road elements such as asphalt pavement, crash barriers, guard-rail, and so on.

This data revealed a general high involvement of the main road network in the occurred rockfalls of this alpine region, a further element that verifies the consistency of the susceptibility map and the chosen approach to compute it. To be noted the relevant frequency of vehicle targets involved (13,6% vehicles sub-classes), followed by people sub-classes target (4,4%) and cattle sub-classes (4,2%), all corresponding to extremely sensitive targets present on roads. Notably, in the observed period (January 2000 - December 2023), national, regional and municipal roads were involved by past rockfall events 213 times, with one or more blocks affecting the road pavement with slight to severe damage to the road elements as guard rail, asphalt and other facilities that led to 26 closures and 35 interruptions of the road, corresponding to about the 4% of the analysed forms. To be noted the involvement of diverse type of vehicles, as cars, trucks or vans (78 times), with an involvement of people in 25 cases. In several cases, also buildings (28.2%) and other anthropic structures and infrastructures (12.5%) as waterworks, power line, parking, and sport facilities have been affected.

Protective measures information extraction

In mountain region, structural barriers are commonly designed and installed along the road network to reduce rockfall impacts and level of risk (Etag 2008). These structural protective measures have been installed over the past decades and include rockfall netting, ring net, mesh drapery system, rockfall barriers, embankments, as well as older system such as the rigid barriers. The evaluation of the involvement of these protective measures by the past rockfalls surely emphasizes the occurrence of this type of phenomenon along major road system, along which such structures are commonly installed, representing a useful element for the analysis of the susceptibility map over time. Additionally, these results confirm the validity of the land planners' choices in placing those barriers. The text mining extraction allowed to primarily identify the kind of protective measure and the frequency of their interaction with past events. Notably the number of observations revealed as the structures most affected are the nets, with a frequency of about 85% including rockfall netting and mesh drapery system, followed by embankment structures (about 10%). These

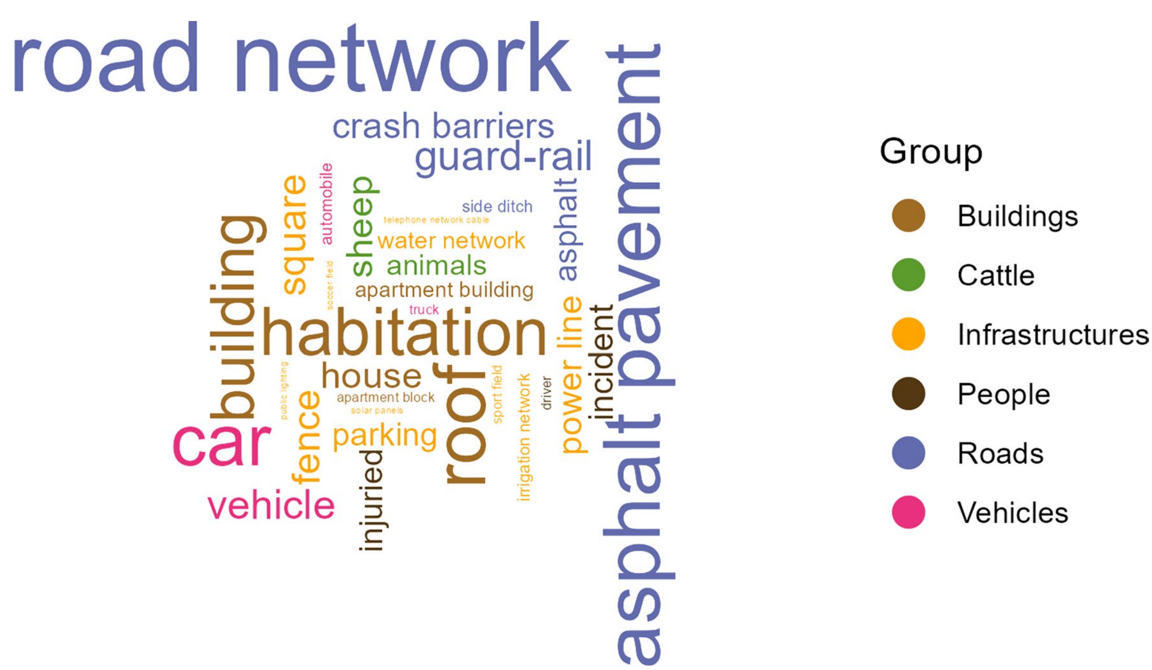

Fig. 6 Word cloud generated after text analysis of the EDFs of the rockfall inventoried in the CD, focusing on the selected keywords referring to the type of target affected, and subdivided in six sub-group reported in the legend

observations confirmed that the inventoried rockfall frequently occurred in sectors with existing structural protective measures, commonly installed in identified, highly exposed and susceptible areas. Analysing the photos collected for each event and reported in the EDF (Fig. 7), we detected that in some cases, the rockfall event has impacted on the rockfall barrier or net, without breaking the structure, while in other cases, the rockfall occur just close or upstream existing structural protective measures.

In addition to structural protective measure, through text mining analysis the protective role of forests was examined. In mountain area, as reported in literature (Fuhr et al. 2015; Moos et al. 2018; Scheidl et al. 2020) forests represent another significant feature in reducing risk related to rockfall events. An extraction of keywords referring to the involvement of trees or plants as computed, observing a frequency of 20.44% of single trees, mainly coniferous, interfering with one more falling block.

Data on the involvement of protective measures may represent a key element for decision-maker for plan a proper management and maintenance of the diverse type of rockfall barriers, as well as of forestry damaged. Further and deeper characterization of the effectiveness of the structures and forestry in protecting the road, could be a great added value for a proper administration of the road network and to schedule the appropriate actions and risk management measures. Such degree of detail could be achieved through the creation of an improved storage of the existing data as a database capable of integrating the vast amount of textual information reachable in each EDF of the regional inventory in a structured and standardized way. These features will allow searching through the dataset by ordinary approach such as database querying thus avoiding incomplete or partial results.

Fig. 7 (**a**) Damage to a rockfall barriers along a municipal road, a block of about 0.7 m^3 hit a vehicle causing serious damage, Cogne (EDF cod. 34933); (**b**) metal rockfall barriers with horizontal elements place in the 90s to protect the regional road S.R.39 destroyed by a rockfall of about 1m^3, Morgex (cod. 36135); (**c**) rockfall event of about 6–8 m^3 causing the rupture of the mesh drapery system and interrupting the municipal road with severe damage to the road pavement, Pontey (cod. 36150); (**d**) rockfall barrier upstream a cheese factory completely removed for 30 m in length by a rockfall of about 30 m^3, Valgrisenche (cod. 36393)

Secondary road system information extraction

Apart from the main road network, in this alpine region, diverse rockfall events also involved forestry roads, footpaths, tourist routes or farm tracks, which represent a very extended (2796.79 km forestry, 5652.67 km footpath) and dense network distributed throughout the entire region. By jointly analysing the road network with both secondary road (i.e., forestry road and farm tracks) and footpaths, text mining analysis revealed a frequency distribution of 31% for main roads, 48.5% forestry road and farm track, and 20.5% of footpath (Fig. 8).

The obtained results pointed out the relevant involvement of the secondary road system and their interaction with the occurred rockfalls, testifying also severe damage to wooden fence, structural protective measures, up to a partial or total occlusion of the path (Fig. 9).

Preliminary frequency assessment of rockfall occurrence

Thanks to the availability of the described dataset, it is also possible to carry out a preliminary check of rockfall frequency and point out areas featuring a higher exposure to these events. Several cases with recurrent rockfall events in the same areas were, in fact, observed. Along the regional road (S.R.) 25, in Arvier municipality, on May 9, 2020 (Fig. 10a) a block of less than 10 m^3 affected the hairpin bends at km 9.100, hitting a passing car and causing damage to car window (Fig. 10b). Three other events occurred in the same place, respectively: on October 30, 2019 (Fig. 10c), consisting in a block of about 0.2–0.3 m^3; on November 11, 2019 (Fig. 10d), consisting in a small rockfall of less than 10m^3 occurred at km 9.100; September 13, 2015 (Fig. 10e), consisting in four blocks of less than 10 m^3 of volume affecting precisely the same area at km 9.100. All these past phenomena stopped on the road pavement without causing severe damage.

Other relevant recurrence was observed along the state highway (S.S.) 26 of the "Piccolo San Bernardo" pass (connecting Italy to France), La Thuile municipality, along the hairpin bends leading to the renowned resort of Pre-Saint Didier. On December 25, 2022 a massive rockfall of 1500 m^3 variably affected the main road, with large block along the hairpin bends and severe damage to the road pavement and structures and infrastructures (e.g., guard rail, power line). In the same area, another event had occurred on October 27, 2018, consisting in a large block that caused severe damage to the S.S.26 and

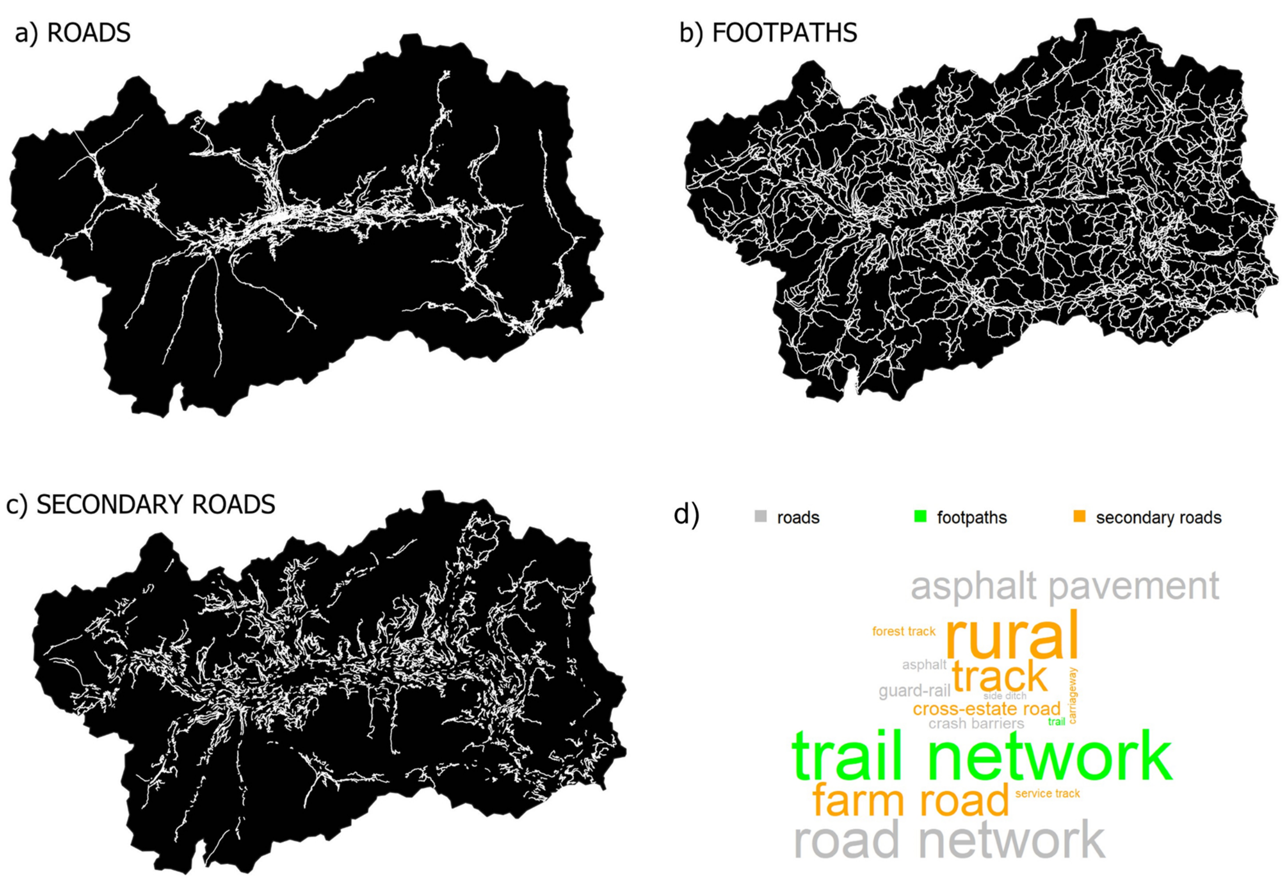

Fig. 8 Spatial distribution of different road types network: (**a**) main roads, (**b**) footpaths and (**c**) secondary roads, and the associated world cloud generated by text analysis, focusing on the selected keywords referring to the three type of roads

Fig. 9 (**a**) Rockfall occurred in correspondence of the footpath to the Fort of Bard, renowned tourist site; most blocks have obstructed part of the path, while a block of 3 m^3 has broken through the parapet on the valley side, reaching the area close to a parking (EDF cod. 36421); (**b**) Wooden fence of a footpath destroyed by a rockfall event that involved the gabions on the valley side, Perloz (cod. 36660); (**c**) Rockfall occurred upstream of the track of a cement factory, with a block of 1.5 m^3 intercepted by the existing rockfall barriers, Courmayeur (cod. 35927); (**d**) small blocks fallen along the "*Strada Romana delle Gallie*", a renowned footpath to the archaeological site of the Roman Arch, Donnas (cod.35863), the event occurred just below the nets and the rockfall barriers

to the protective works along the road. Although briefly described, an event in the same area that occurred on July 11, 1993 is described in the CD as a rockfall of about 3 m^3 that caused damage to the road pavement.

Through a systematic extraction and analysis of volume and occurrence time of each event, the volume/frequency relationship could be computed, obtaining a unique and valuable dataset of more than two thousands of rockfalls occurred over more than eight hundreds years. As demonstrated in Fei et al. 2023, this approach, applied at slope scale, creates the basis for the evaluation of rockfall event return period, and should be applied to regional scale context, as the Aosta Valley one, as a tool to identify more affected and damaged areas with the

Fig. 10 Comparison between a post-February rockfall occurred in Arvier municipality, and the past rockfall events collected in the CD: (**a**) May 9, 2020 event, a block hit a passing car hitting the car window (**b**) at km 9.100 of the S.R.25 (cod. 34969); (**c**) October 30, 2019 event, a block reached the S.R.25 without record severe damage (cod. 34654); (**d**) November 6, 2019 event, a small rockfall occurred along the S.R.25 at km9.100, stopped on the road pavement (cod. 34657); (**e**) September 13, 2015, four blocks affecting precisely the same area at km 9.10; (**f**) location of all the rockfall events occurred in the same stretch of the hairpin bends of the S.R.25 at km 9.100, in red the post-February event (May 9, 2020) and in black the previous events recorded: 1 = September 13, 2015; 2 = October 30, 2019; 3 = November 6, 2019 event

aim of planning appropriate protection and mitigation measures.

The availability of an events database is a key factor in the management of rockfall (or other slope failure phenomena) hazards, leading to actions like susceptibility maps computation or land planning. Governments, local or national, with country sectors endangered by those events, like the study area, should implement procedures to compile an inventory to better cope with rockfalls and their risk assessment. Example of those archives can be found worldwide at national (Rupp and Damm 2020) or site specific (Perret et al. 2006) levels; alternatively, they can be compiled with the purpose of risk assessment along an endangered infrastructure (Regmi et al. 2016). When unavailable the databases should be carefully designed to correctly store all the information and data collected during field surveys and documents collection. Besides, the data collection phase should be meticulously planned to avoid surveying uncertain data, e.g. location of the detachment zone being confused with impact zone. Lastly the database should be implemented in an Open Source architecture so as to facilitate data portability and exchangeability.

Conclusions

Landslide susceptibility mapping has proven to be a reliable way to preliminary identify area featuring higher proneness to slope instability, and usable as a tool in land use and engineering interventions planning. The usability of such maps over time and their reliability therefore plays a key role to address the needs of technicians and decision-makers in remedial works designing to protect man-made structures and infrastructure. The *a-posteriori* analysis of the rockfall susceptibility map along the main road network of the Aosta Valley allowed us to verify its consistency over time. From the dual-analysis, in GIS-environment and through text-mining application, on information collected in the regional landslide inventory, an effective exploitation of both geocoded and documentary data was ensured for the assessment of map reliability. Exploiting the in detail and continuously updated regional landslide inventory of this alpine region, we derived a reliable dataset of the event occurred after three years from the production of the map. Testing the prior output mapping, based on the AHP approach, by comparing the four main classes of susceptibility zonation to the rockfall events occurred after the publication of the map (post-February 2020), relevant correspondences have been detected, with about 80% of the latest events located in areas with high and very high susceptibility. These results indicate that the recent events consistently match the previous susceptibility map, pointing out the predictive capability of this product over time. Moreover, the text mining approach provided additional information to better characterize the impact of the analysed phenomena on the studied area by identifying the most affected targets and the related degree of damage and set up a preliminary frequency analysis. The analysis allowed us to point out that, other than main roads, also other infrastructures are affected by rockfalls such as farm and forestry roads, a strategic network in such an alpine area characterized by a remarkable presence of farming and forestry practices, and footpaths that covers a notable percentage of regional territory and are utilized by thousands of tourists per year. Overall, susceptibility maps certainly proven to be a useful tool for land planning, however their validity is, currently, limited to the date of elaboration; a feasible improvement is the validity check with *a-posteriori* approach, like the one proposed, to verify their consistency over time and their usability for risk management purposes.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40677-024-00300-w.

Supplementary Material 1

Acknowledgements

The staff of "Assessorato Agricoltura e Risorse naturali, Dipartimento risorse naturali e Corpo forestale, Struttura foreste e sentieristica" for providing data and technical support to the analysis in the farm and forestry sector.

Author contributions

Conceptualization: MC, DGo; Data curation: MC, DGo, DC; Methodology: MC, DGo; Data analysis: MC, DGo; Writing-original draft preparation: MC, DGo, DC; Reviewing and editing draft: MC, DGo, DC, DGi, DB. All authors have read and approved the final manuscript.

Funding

Not applicable.

Data availability

No datasets were generated or analysed during the current study.

Declarations

Competing interests

The authors declare no competing interests.

Received: 8 April 2024 / Accepted: 16 October 2024
Published online: 31 October 2024

References

Alvioli M, Santangelo M, Fiorucci F, Cardinali M, Marchesini I, Reichenbach P, Rossi M, Guzzetti F, Peruccacci S (2021) Rockfall susceptibility and network-ranked susceptibility along the Italian railway. Engineering Geology 293. Elsevier: 106301. https://doi.org/10.1016/J.ENGGEO.2021.106301

Bistacchi A, Dal Piaz G, Massironi M, Zattin M, Balestrieri M (2001) The aosta–ranzola extensional fault system and oligocene–Present evolution of the austroalpine–penninic wedge in the northwestern Alps. Int J Earth Sci 90:654–667. https://doi.org/10.1007/s005310000178

Brabb EE (1984) Innovative approaches to landslide hazard and risk mapping. publisher not identified

Carrara A, Cardinali M, Detti R, Guzzetti F, Pasqui V, Reichenbach P (1991) GIS techniques and statistical models in evaluating landslide hazard. Earth Surface Processes and Landforms 16. John Wiley & Sons, Ltd: 427–445. https://doi.org/10.1002/ESP.3290160505
Carraro F, Giardino M (2004) Quaternary glaciations in the western Italian Alps-a review. Elsevier
Centro Funzionale (2020) Regione Autonoma Valle d'Aosta. Catasto Dissesti
Cignetti M, Godone D, Bertolo D, Paganone M, Thuegaz P, Giordan D (2020) Rockfall susceptibility along the regional road network of Aosta Valley Region (northwestern Italy). Journal of Maps 17(3). Taylor & Francis: 54–64. https://doi.org/10.1080/17445647.2020.1850534
Clemente J, Ander JA, Uriarte D, Spizzichino F, Faccini, Tomás, Morales (2023) Rockfall hazard mitigation in coastal environments using dune protection: A nature-based solution case on Barinatxe beach (Basque Coast, northern Spain). *Engineering Geology* 314. Elsevier: 107014. https://doi.org/10.1016/J.ENGGEO.2023.107014
Core Team RD (2022) R. R: A Language and Environment for Statistical Computing. Edited by R Development Core Team. Vol. 1. R Foundation for Statistical Computing. R Foundation for Statistical Computing. https://doi.org/10.1007/978-3-540-74686-7
Doan Duong, Hai QD, Dang KD, Huu EF, Aqidawati W, Sutopo MH et al (2016) Newspaper archives + text mining = rich sources of historical geo-spatial data. IOP Conference Series: Earth and Environmental Science 34. IOP Publishing: 012043. https://doi.org/10.1088/1755-1315/34/1/012043
Etag EOTA (2008) 027-Guideline for the European Technical Approval of Falling Rock Protection Kits. EOTA: Brusseles, Belgium
Fei, Li M, Jaboyedoff A, Guerin François, Noël D, Bertolo M-H, Derron P, Thuegaz F, Troilo, and Ludovic Ravanel (2023). Assessing the rock failure return period on an unstable Alpine rock wall based on volume-frequency relationships: The Brenva Spur (3916 m asl, Aosta Valley, Italy). Engineering Geology 323. Elsevier: 107239
Fuhr M, Bourrier F, Cordonnier T (2015) Protection against rockfall along a maturity gradient in mountain forests. Forest Ecology and Management 354. Elsevier: 224–231. https://doi.org/10.1016/j.foreco.2015.06.012
Furlani S, Ninfo A (2015) Is the present the key to the future? Earth-Science Reviews 142. Elsevier: 38–46. https://doi.org/10.1016/J.EARSCIREV.2014.12.005
Gianotti F, Forno MG, Ivy-Ochs S, Peter WK (2008) New chronological and stratigraphical data on the Ivrea amphitheatre (Piedmont, NW Italy). Quatern Int 190:123–135. https://doi.org/10.1016/j.quaint.2008.03.001
Giardino M, Ratto SM (2007) Analisi del dissesto da frana in Valle d'Aosta
Giordan D, Cignetti M, Godone D, Peruccacci S, Raso E, Pepe G, Calcaterra D et al (2020) A new procedure for an effective management of geo-hydrological risks across the Sentiero Verde-Azzurro Trail, Cinque Terre National Park, Liguria (North-Western Italy). Sustainability 12. MDPI AG: 561. https://doi.org/10.3390/su12020561
Giordan D, Cignetti M, Godone D, Bertolo D, Paganone M (2021) Definition of an Operative Methodology for the Management of Rockfalls along with the Road Network. Sustainability 2021, Vol. 13, Page 7669 13. Multidisciplinary Digital Publishing Institute: 7669. https://doi.org/10.3390/SU13147669
Guzzetti F, Reichenbach P, and Silvia Ghigi (2004) Rockfall Hazard and Risk Assessment along a Transportation Corridor in the Nera Valley, Central Italy. Environ Manage 34:191–208. https://doi.org/10.1007/s00267-003-0021-6
Guzzetti F, Carrara A, Cardinali M, Reichenbach P (1999) Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. Geomorphology 31:181–216. https://doi.org/10.1016/S0169-555X(99)00078-1
Guzzetti F, Crosta G, Detti R, Agliardi F (2002) STONE: a computer program for the three-dimensional simulation of rock-falls. Comput Geosci 28:1079–1093
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
Hungr O, Evans SG, Hazzard J (1999) Magnitude and frequency of rock falls and rock slides along the main transportation corridors of southwestern British Columbia. Canadian Geotechnical Journal 36. NRC Research Press: 224–238
ISPRA Ambiente (2012) CARG Project
ISTAT (2023) Il Censimento permanente della popolazione in Valle d'Aosta/Vallée d'Aoste in Censimento permanente popolazione e abitazioni.
Moos C, Fehlmann M, Trappmann D, Stoffel M, Dorren L (2018) Integrating the mitigating effect of forests into quantitative rockfall risk analysis – Two case studies in Switzerland. International Journal of Disaster Risk Reduction 32. Elsevier Ltd: 55–74. https://doi.org/10.1016/j.ijdrr.2017.09.036
Palma B, Parise M, Reichenbach P, Guzzetti F (2012) Rockfall hazard assessment along a road in the Sorrento Peninsula, Campania, southern Italy. Natural Hazards 61. Springer: 187–201. https://doi.org/10.1007/s11069-011-9899-0
Palomba M, Giardino M, Ratto S, Pogliotti P (2015) Analysis of Factors Controlling Landslide Susceptibility in the Aosta Valley (NW Italy): Relationship to Climatic and Environmental Changes. In Engineering Geology for Society and Territory - Volume 1: Climate Change and Engineering Geology, 435–438. Springer International Publishing. https://doi.org/10.1007/978-3-319-09300-0_83
Perret S, Stoffel M, and Hans Kienholz (2006) Spatial and temporal rockfall activity in a forest stand in the Swiss Prealps—A dendrogeomorphological case study. Geomorphology 74:219–231. https://doi.org/10.1016/j.geomorph.2005.08.009
Powell G, and others (2002) Discussion'landslide Risk Management concepts and guidelines'. Australian Geomechanics: J News Australian Geomech Soc 37:45
Raetzo H, Lateltin O, Bollinger D, Tripet J (2002) Hazard assessment in Switzerland – codes of Practice for mass movements. Bull Eng Geol Environ 61:263–268. https://doi.org/10.1007/s10064-002-0163-4
Ratto S, Bonetto F, Comoglio C (2003) The October 2000 flooding in Valle d'Aosta (Italy): Event description and land planning measures for the risk mitigation. International Journal of River Basin Management 1: 105–116. https://doi.org/10.1080/15715124.2003.9635197
Regione Autonoma Valle d'Aosta (2020) Land Cover RAVA 2020, Sistema delle Conoscenze Territoriali. https://geoportale.regione.vda.it/download/carta-copertura-suolo/. Accessed 31 Oct 2023
Regmi A, Deep P, Cui MR, Dhital, Zou Q (2016) Rock fall hazard and risk assessment along Araniko Highway, Central Nepal Himalaya. Environmental Earth Sciences 75. Springer: 1–20
Reichenbach P, Rossi M, Malamud BD, Mihir M, Guzzetti F (2018) A review of statistically-based landslide susceptibility models. Earth-Science Reviews. Elsevier B.V. https://doi.org/10.1016/j.earscirev.2018.03.001
Rupp S, Damm B (2020) A national rockfall dataset as a tool for analysing the spatial and temporal rockfall occurrence in Germany. Earth Surface Processes and Landforms 45. Wiley Online Library: 1528–1538
Saaty TL (1980) Analytic hierarchy process. Wiley Online Library
Salloum SA, Al-Emran M, Monem AA, Shaalan K (2018) Using text mining techniques for extracting information from research articles. Studies in Computational Intelligence 740. Springer Verlag: 373–397. https://doi.org/10.1007/978-3-319-67056-0_18/FIGURES/17
Scheidl C, Heiser M, Vospernik S, Lauss E, Perzl F, Kofler A, Kleemayr K et al (2020) Assessing the protective role of alpine forests against rockfall at regional scale. Eur J for Res 139 Springer Sci Bus Media Deutschland GmbH 969–980. https://doi.org/10.1007/S10342-020-01299-Z/FIGURES/5
Shirzadi A, Saro L, Joo OH, Chapi K (2012) A GIS-based logistic regression model in rock-fall susceptibility mapping along a mountainous road: Salavat Abad case study, Kurdistan, Iran. Natural Hazards 64. Springer: 1639–1656. https://doi.org/10.1007/s11069-012-0321-3
Trigila A, Iadanza C, Spizzichino D (2008) IFFI Project (Italian landslide inventory) and risk assessment. In Proceedings of the First World Landslide Forum, Tokyo, Japan, 18–21November 2008, pp 18–21
Varnes DJ (1984) Landslide hazard zonation: A review of principles and practice. Natural Hazards
Yagoub MM, Aishah A, Alsereidi, Elfadil A, Mohamed P, Periyasamy R, Alameri S, Aldarmaki, and Yaqein Alhashmi (2020). Newspapers as a validation proxy for GIS modeling in Fujairah, United Arab Emirates: identifying flood-prone areas. Natural Hazards 104. Springer: 111–141

Publisher's note