

An inventory-driven rock glacier status model (intact vs. relict) for South Tyrol, Eastern Italian Alps

Christian Kofler^{a,b}, Stefan Steger^a, Volkmar Mair^c, Marc Zebisch^a, Francesco Comiti^b, Stefan Schneiderbauer^{a,d}

^a Institute for Earth Observation, Eurac Research, 39100 Bozen/Bolzano, Italy

^b Faculty of Science and Technology, Free University of Bozen-Bolzano, 39100 Bozen/Bolzano, Italy

^c Office for Geology and Building Materials Testing, 39053 Kardaun/Cardano, Italy

^d UNU-EHS, GLOMOS Programme, 39100 Bozen/Bolzano, Italy

ARTICLE INFO

Article history:

Received 4 April 2019

Received in revised form 16 August 2019

Accepted 24 September 2019

Available online 28 October 2019

Keywords:

Machine learning

Probabilistic modelling

Permafrost

Classification algorithms

Rock glacier inventory

Rock glacier activity

ABSTRACT

Ice presence in rock glaciers is a topic that is likely to gain importance in the future due to the expected decrease in water supply from glaciers and the increase of mass movements originating in periglacial areas. This makes it important to have at ones disposal inventories with complete information on the state of rock glaciers. This study presents a method to overcome incomplete information on the status of rock glaciers (i.e. intact vs. relict) recorded in regional scale inventories. The proposed data-driven modelling framework can be used to estimate the likelihood that rock glaciers contain frozen material. Potential predictor variables related to topography, environmental controls or the rock glacier appearance were derived from a digital terrain model (DTM), satellite data and gathered from existing data sets. An initial exploratory data analysis supported the heuristic selection of predictor variables. Three classification algorithms, namely logistic regression (GLM), support vector machine (SVM) and random forest (RF), were trained on the basis of the available information on the status of rock glaciers within the territory of South Tyrol (Eastern Italian Alps). The resulting classification rules led to assign a binary label – intact or relict – to 235 unclassified rock glaciers present in the inventory. All models were validated quantitatively on spatially-independent test samples (spatial cross validation) and achieved highly satisfactory performance scores. Hereby, the less flexible statistically-based classifier (GLM) performed slightly better than the more flexible machine learning algorithms (SVM and RF). Spatial permutation-based variable importance assessment revealed that elevation and vegetation cover (based on NDVI) were the most relevant predictors. For more than 80% of the unclassified rock glaciers, all of the three models agreed on the spatially predicted rock glacier status. Only for a minor portion (12.3%), one model differed from the remaining two.

© 2019 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Rock glaciers are cryogenic landforms of cold mountain landscapes that are formed by downslope creeping, perennially frozen debris (Haeberli et al., 2006). Presently ice-saturated rock glaciers are the visible expression of active “sediment conveyor systems” from sediment source areas towards their front (Müller et al., 2014). In case of an existing connection to the channel network, rock glaciers can also be considered a sediment source (e.g. Kummert and Delaloye, 2018), and the supply of loose material from rock glaciers into torrents can therefore be relevant in terms of debris flow hazards (e.g. Kummert et al., 2017). Besides their role as controlling landform of sediment cascades in cold mountain environments (e.g. Müller et al., 2014), ice-rich rock glaciers act as

important fresh water storages and sources, especially in semi-arid and arid mountainous areas. As highlighted by Brenning (2003), the amount of water stored in intact rock glaciers might even exceed water volumes stored in glaciers within catchments of the Chilean Andes. Recently, Jones et al. (2018a) emphasized the high resilience against climatic warming of rock glaciers compared to glaciers, as well as their future relevance for freshwater supply. In contrast, ice-free rock glaciers are evidence of paleo-climatic conditions (e.g. Cossart et al., 2010; Frauenfelder and Kääb, 2000)(cf. Chapter 2 for a detailed elaboration of the study object).

Information on the spatial distribution, status and geometry of rock glaciers is usually stored in dedicated GIS-based inventories (e.g. Colucci et al., 2016; Cremonese et al., 2011; Jones et al., 2018b; Scotti et al., 2013). However, time and cost constraints do

not permit to assign the status to each rock glacier in regional scale inventories based on in-situ evidence. Therefore, the determination of the status of an individual rock glacier is mostly carried out based on expert knowledge with support of aerial imagery, digital terrain models (DTMs), field investigations and other auxiliary material. This is a reasonable approach but makes the final inventory prone to biases related to subjective estimations of different operators in charge of defining the status of rock glaciers, as shown by [Brardinoni et al. \(2019\)](#). In addition, it hampers comparability among different inventories.

Remote sensing technologies can provide useful data to investigate geomorphic processes in cold mountain areas (e.g. [Barboux et al., 2014](#); [Delaloye et al., 2008](#); [Prantl et al., 2018](#)). [Bertone et al. \(2019\)](#) have recently shown that freely available synthetic aperture radar (SAR) data can offer an opportunity to complement the well-established expert-based evaluation of rock glacier status with more quantitative methods.

Rock glacier inventories have been widely utilized as a data source within multiple variable models to study environmental controls on rock glaciers or mountain permafrost distribution (e.g. [Boeckli et al., 2012](#); [Sattler et al., 2016](#); [Azócar et al., 2017](#); [Brenning et al., 2007](#)). The logistic regression model was among the first statistical classifiers applied to study the relationship between rock glacier occurrence and morphometric or topoclimatic control variables. Such models are able to reproduce complex relationships among rock glaciers and their determining conditions such as an elevation-dependent incoming solar radiation ([Brenning and Trombotto, 2006](#)). Later, also machine learning models were utilized for rock glacier detection or permafrost distribution modelling (e.g. [Brenning, 2009](#); [Deluigi et al., 2017](#)).

The objective of this paper is to present data-driven models to estimate the status of rock glaciers at the regional scale by integrating various geospatial datasets, such as high-resolution terrain data and satellite data products. Additionally, the paper aims at exploring the model results in terms of empirical relationships between rock glacier status and the controlling environmental variables. The pre-requisites to conduct this study were (i) the availability of a dedicated inventory, which contains mapped rock glaciers whose status is unknown, and (ii) a training data set with information on the status of rock glaciers. Spatial intrinsic and environmental features related to rock glaciers are integrated into multiple variable models to investigate their ability to compute the relative probability of rock glaciers to contain frozen material or not (i.e. to be intact or relict). A well-established probabilistic classifier (logistic regression) was compared with two flexible machine-learning algorithms (support vector machine and random forest).

The structure of the paper consists of an initial literature review and exploratory data analysis that supported the identification of appropriate controlling variables of rock glacier status (cf. Sections 2 and 4.2–4.4.). The subsequent construction of three multiple-variable models was accompanied by an in-depth model evaluation based on spatial cross validation and geomorphological plausibility (cf. Section 4.5–4.8.). The final models were applied to assign the status to 235 unclassified rock glaciers present in the inventory of South Tyrol (Northern Italy). Finally, the modelling results were analyzed in terms of their quantitative response (i.e. response curves for selected variables), variable importance and spatial prediction agreement (i.e. inventory vs. models; discrepancies among the models).

2. Rock glacier definition, origin and controlling factors

The presence of rock glaciers in high mountain landscapes is a result of topographic, climatic and geological controls, which determine their development and spatial distribution. The visual

appearance of rock glaciers has been described as tongue- or lobate shaped, sometimes also spatulate-shaped with an uneven surface characterized by the presence of longitudinal or transversal ridges and furrows. Another distinct morphological feature are the terminal front and lateral slopes whose angle is near or even beyond the internal angle of friction of the composing material. Ice-rich rock glaciers can be identified by a rather convex surface as well as a general absence of vegetation. If inactive, their marginal slopes at the front and the sides are invaded by some vegetation patches ([Giardino and Vitek, 1988](#); [Haeberli, 1985](#); [Potter, 1972](#); [Wahrhaftig and Cox, 1959](#)). On the other hand, relict rock glaciers, can – depending upon their location – also be completely vegetation-covered, as shown by [Colucci et al. \(2016\)](#) (see Fig. 1).

In general, rock glaciers are built up by poorly-sorted angular rock debris and ice, but the superficial and the internal composition may greatly differ. Indeed, the surface of rock glaciers is dominated by pebbly to bouldery clasts ([Ikeda and Matsuoka, 2006](#)), whereas, below a certain depth, the grain size range increases and fine material tends to dominate ([Krainer et al., 2015b](#)). Below the seasonally unfrozen active layer, ice can be present as pure ice, ice in the voids between particles or as ice lenses ([Arenson, 2002](#); [Jones et al., 2019](#); [Krainer et al., 2015b](#)). The deformation of this internal ice/rock mixture enables the steady state creep of rock glaciers. The origin of the composing sediment may be either rock fall talus or glacially eroded material. [Scotti et al. \(2013\)](#) state that between the end of the Last Glacial Maximum (LGM) and the beginning of the Holocene, rock glaciers in the Lombardy region (Central Eastern Alps) were mainly fed by talus, while till-rock glaciers originated mostly during the Little Ice Age (LIA). In their inventory, this latter type is predominantly intact, while the former type is mostly relict. [Lilleøren and Etzelmüller \(2011\)](#), too, support the assumption that rock glaciers formed in connection with glacial activity are frequently younger than talus-derived rock glaciers. For their study sites in Southern and Northern Norway, they conclude that talus-derived rock glaciers can be seen as relict landform, representing a former dry, periglacial climate.

Statistical analyses of the rock glacier inventories of South Tyrol ([Bollmann et al., 2012](#)) and of its neighboring Trentino Province ([Seppi et al., 2012](#)) highlighted that both the total percentage of rock glaciers as well as the share of intact rock glaciers is low in mountain ranges that are dominated by carbonate and dolomitic rocks. In contrast, mountain ranges formed by metamorphic rocks generally show a higher number of rock glaciers, particularly of intact ones. Relations between lithology and rock glacier presence were obtained also by studies conducted in the Swiss and Austrian Alps, as well as in Idaho (USA) ([Johnson et al., 2007](#); [Kellerer-Pirklbauer, 2007](#); [Matsuoka and Ikeda, 2001](#)).

The origin of ice in rock glaciers is an intensively debated topic within the scientific community. Two opposing theories, namely the “permafrost model” and the “glacial ice core model” have been proposed. The former one assumes a nonglacial setting where the internal ice of the rock glacier derives from meltwater, groundwater or buried (avalanche) snow and past or present glaciation is here not seen as a factor for their genesis ([Haeberli et al., 2006](#); [Jones et al., 2019](#); [Wahrhaftig and Cox, 1959](#)). On the other side, the promoters of the glacial ice core model accept permafrost as a forming condition for rock glaciers but do not state this to be a mandatory requirement. This model assumes a thin core of glacial ice that is protected by a layer of weathered rock debris. The transition from glaciers to rock glaciers is assumed to be a continuous process with varying degrees of ice/rock debris ratios ([Johnson, 1980](#); [Martin and Whalley, 1987](#); [Potter, 1972](#); [Whalley and Martin, 1992](#)). Recently, this conceptual model has been supported by geophysical and remote sensing observations as well as computational models, which consider a rock glacier as the end-member state in a

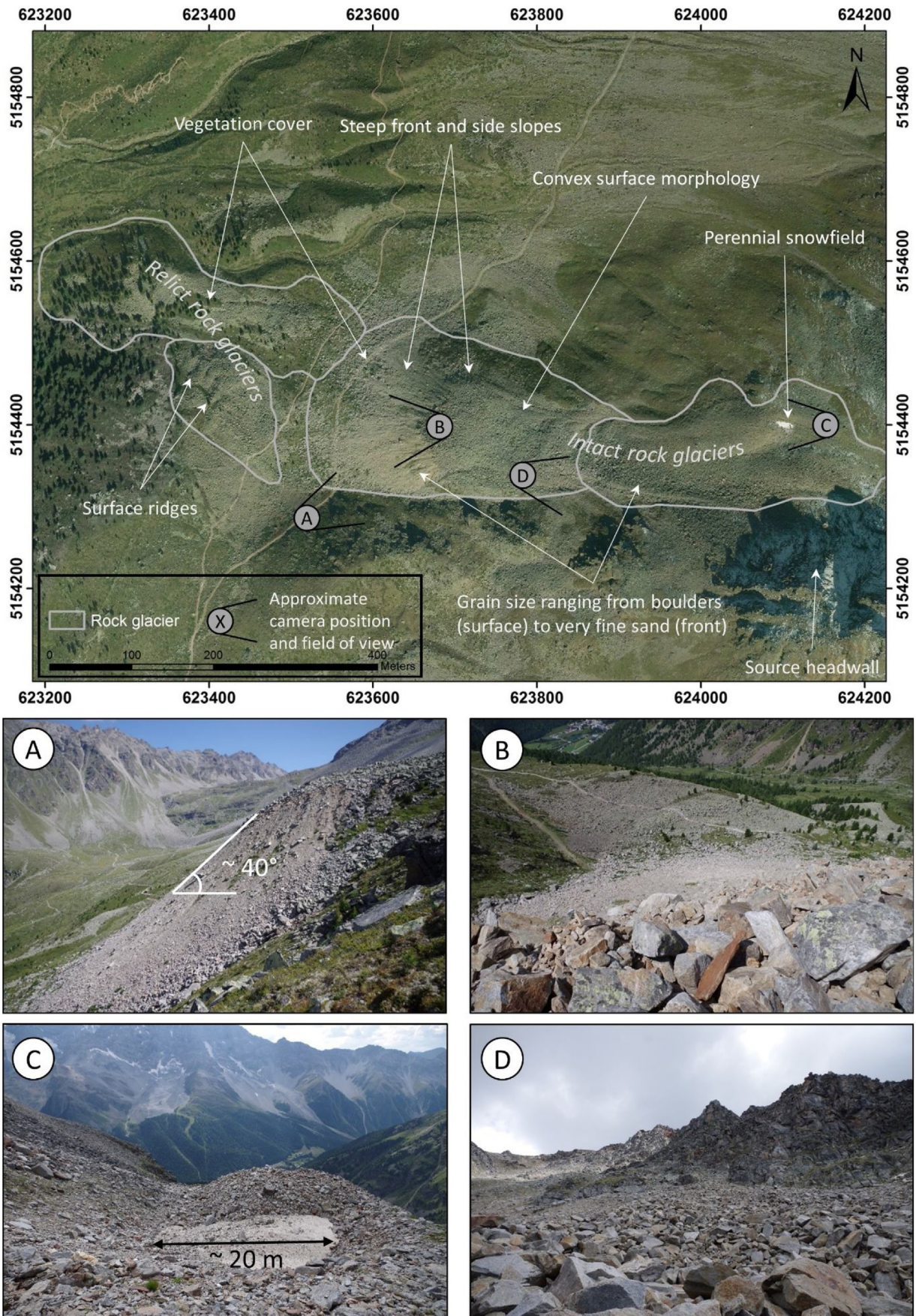


Fig. 1. Morphological characteristics of rock glaciers. The map on top shows intact and relict tongue-shaped rock glaciers located in Zay Valley - South Tyrol, at $46^{\circ} 31.911' \text{N}$, $10^{\circ} 36.763' \text{E}$ (cf. Fig. 6). A: View from below on the rock glacier front; B: View from the edge of the intact rock glacier surface on the relict rock glaciers; C: Perennial snowfield;

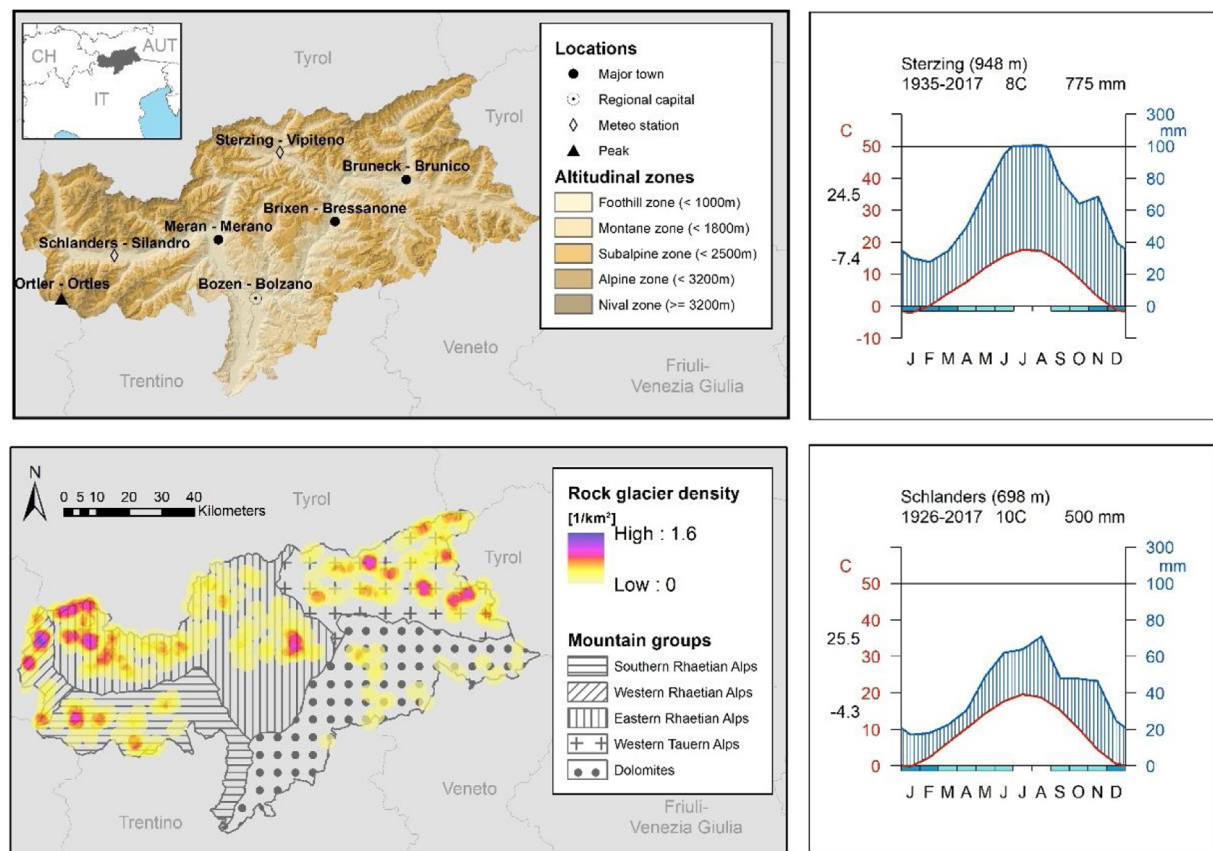


Fig. 2. Overview of the study area and altitudinal zonation (upper left), the subdivision of the main mountain groups according to the SOIUSA classification of Marazzi (2005) and the density of rock glaciers in South Tyrol (lower left; created based on the point density of rock glaciers stored in the inventory of South Tyrol, cf. Section 4.2.). The position of the two meteorological stations (upper and lower right) is shown in the upper left.

glacier – debris covered glacier – rock glacier continuum (Anderson et al., 2018; Monnier et al., 2011; Monnier and Kinnard, 2015). A recent work of Knight et al. (2019) adds a third model, namely the “paraglacial model”, which assumes rock glacier formation as a consequence of increased rock fall activity due to slope relaxation processes after deglaciation.

Berthling (2011) developed a conceptual solution to the debate on the origin of rock glaciers by assigning them neither entirely to the periglacial nor to the glacial realm. Instead he proposed to accept permafrost conditions as a determining spatiotemporal control on the long-term development and creep rate. This proposal follows the concept of “cryo-conditioning” of periglacial landforms proposed by Berthling and Etzelmüller (2011), which emphasizes cryotic surface and subsurface thermal conditions as the common control for the development of cold-region landscapes. For regional-scaled studies, the mean annual air temperature (MAAT) is a commonly used variable to approximate these thermal controls (Boeckli et al., 2012; Gruber, 2012).

The presence of rock glaciers is further determined by the characteristics of the local topography. They are commonly found below the environmental equilibrium line altitude (ELA), which marks the balance between accumulation and ablation of ice on glaciers. Consequently, their presence and status is determined by site characteristics that allow ice to persist throughout the year

(Anderson et al., 2018). Additionally, the usual steep topographic setting around rock glaciers favors the supply of ice and rock debris by snow avalanches, which may feed the rock glacier directly or indirectly through perennial snowfields. In this context, Humlum et al. (2007) introduced the term “avalanche-derived rock glacier” by presenting a conceptual model of snow avalanches providing rock debris and ice to the rock glacier.

Another topography-related factor influencing the presence and distribution of rock glaciers is solar radiation (e.g. Blöthe et al., 2018), and indeed it is often the main local-scale energy source that determines the presence of permafrost conditions (Hoelzle, 1992). Even at the lower altitudes, where MAAT is positive, permafrost can exist if radiation is strongly reduced by topography (Hoelzle and Haeblerli, 1995).

3. Study area

The regional territory of South Tyrol (Eastern Italian Alps, Fig. 2) covers an area of approximately 7,400 km² and its altitude ranges from 194 m to 3,893 m a.s.l. Due to its inner-alpine location, South Tyrol is a relatively dry region with lower precipitation rates compared to areas located at the outer parts of the Alps. Within South Tyrol, the western part belonging to the Southern and Eastern Rhaetian Alps is drier than the central and eastern part due to shad-

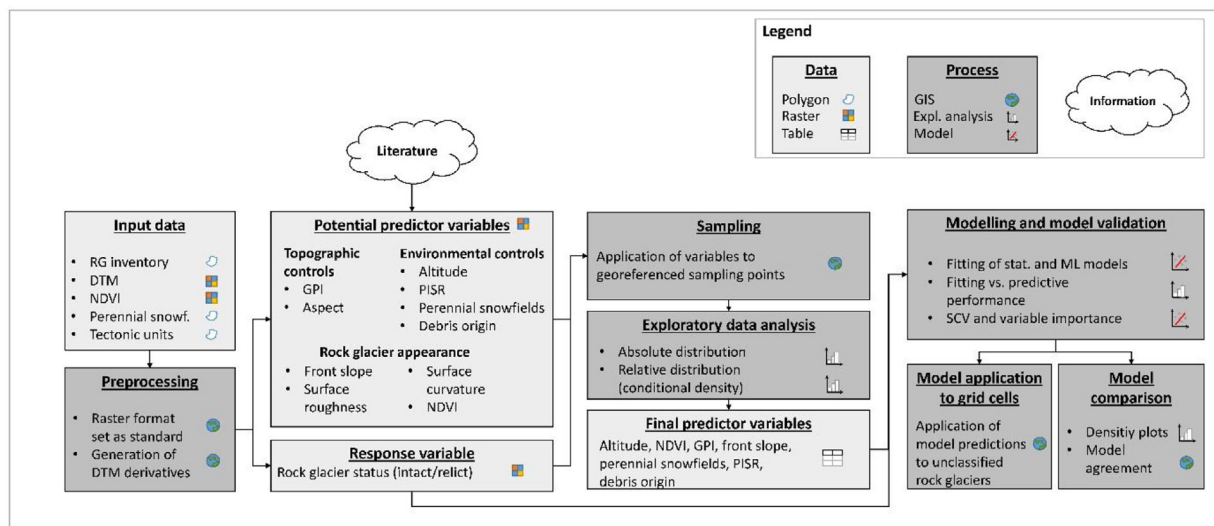


Fig. 3. Flowchart presenting the applied and generated datasets and processing steps. Abbreviations (if not defined before): RG = rock glacier; GPI = geomorphic protection index; PISR = potential incoming solar radiation; NDVI = normalized difference vegetation index SCV = spatial cross validation; DTM = Digital Terrain Model; GIS = Geographic Information System.

ing effects of these mountain groups. In between, one of the driest valleys of the Alps – the Vinschgau/Venosta Valley – is located. This East-West extending valley is characterized by a continental alpine precipitation regime (Della Chiesa et al., 2014). As it can be seen on the climate diagram in Fig. 2, a mean annual precipitation (MAP) of about 500 mm is measured in Schlanders/Silandro, a town in the middle Vinschgau valley bottom. In contrast, a MAP around 775 mm is measured in Sterzing/Vipiteno, located in the upper Eisack/Isarco valley (N-S alignment). Across the whole South Tyrol, precipitation increase with altitude, with relatively important elevation gradients (e.g. Mergili and Kerschner, 2015).

As to geology, the dominant lithologies in the western part of South Tyrol are metamorphic rocks such as schists and gneisses, whereas the south-eastern part is characterized by limestones and dolomites. Igneous rocks are also relevant in some valleys. The spatial distribution of rock glaciers (Fig. 2) shows that their concentration is highest in the NE and W South Tyrol, mainly on the main Alpine ridge (so called Rhaetian and Tauern Alps).

4. Materials and methods

4.1. Workflow description and software used

The methodological workflow of this study (Fig. 3) comprises the identification, collection and preprocessing of input data and the selection of potential predictor variables (Section 4.2. – 4.4.). The exploratory data analysis builds the basis for the heuristic input data selection (Section 4.5.). Three multiple variable classification algorithms were then quantitatively validated within a spatial cross validation framework while also assessing spatial variable importance (Section 4.6. – 4.7.). The final classification rules were transferred to each rock glacier listed in the inventory of South Tyrol (Section 4.8.). Response curves for selected predictors and a confrontation of predicted activity states (i.e. model agreement) provided insights into the quantitative model responses.

All statistical and machine-learning analyses were performed with the front-end software “RStudio” of the open-source statistical computing package “R”, version 3.5.0 (R Core Team, 2017). Some tasks could be fulfilled with the preinstalled package “stats” e.g. the exploratory data analysis and the modelling part based on logistic regression. The support vector machine model (SVM) is based on the package “kernlab” (Karatzoglou et al., 2004), and

the random forest classifier was built through the “randomForest” package (Liaw and Wiener, 2002). The packages “sperrorest” (Brenning, 2012) and “raster” (Hijmans, 2017) were used to validate the model and to transfer it to grid cells. The computation of DTM-derivatives as well as the preparation of the necessary input maps for the analysis was done by the open-source GIS “SAGA GIS”, version 6.3.0 (Conrad et al., 2015) and the associated R-package “RSAGA” (Brenning, 2008). The layout of the maps was designed with the commercial GIS “ArcGIS”, version 10.4 (ESRI, 2016).

4.2. Data

The main data source was the inventory of rock glaciers for South Tyrol. It contains their spatial position as well as the training data on their status, which were necessary for this study. This dataset was compiled by a number of researchers within the project PROALP (funded by the Autonomous Province of Bolzano-South Tyrol), a joint initiative between academic, public and private institutions as described in Bollmann et al. (2012). A total number of 1779 rock glaciers are in the inventory, of which 240 are listed as active, 59 as inactive and 1245 as relict. The status of 235 rock glaciers was not defined yet. The individual rock glacier status was assigned by interpreting geological maps, aerial photographs and the hillshade of the available high resolution (2.5 m) Light Detection and Ranging (LiDAR) digital terrain model (DTM). As reported by the authors, space-borne multi-temporal radar interferometry (MT-InSAR) was applied for single rock glaciers to deduce the status from the kinematic behavior. Besides the status, also information on the debris origin (till/talus), the largest identified block diameter (larger/smaller than 1 m), length and width, altitude and slope aspect of each individual rock glacier is available. The data set is freely available in Shapefile format on the WebGIS platform of the provincial administration (<http://geokatalog.buergernetz.bz.it/geokatalog>¹).

Many of the applied predictor variables in this study’s modelling setup (cf. Chapter 4.4.) were derived from the same DTM mentioned above. It covers the entire province and refers to surveys carried out in 2005.

¹ Access Date: 19.03.2018

In addition, semi-automatically mapped data on perennial snowfields were considered in the present study. The dataset was created by Zischg et al. (2012) on the basis of aerial images acquired in late summer (1996, 1999 and 2003). In order to obtain a map that represents a proxy variable for the temporal and spatial persistence of perennial snow patches, the areas covered with snow in all the three years were labelled as such.

Information on the vegetation cover was approximated with the normalized difference vegetation index (NDVI). The NDVI derives from optical remote sensing imagery and can be used as indicator for vegetation presence (Rouse et al., 1974). The NDVI maps used in this study are a product derived from openly accessible satellite scenes of the European Space Agency's (ESA) Sentinel 2 A/B mission. The processing chain included a correction of raw images (atmospheric correction, removal of clouds, snow and artificial surfaces) and the calculation of NDVI values with Band 03 (red) and Band 08 (near infrared) (see Rossi et al., 2019 for a detailed description). The final product is presented with a grid resolution of 10 m. Only derivatives of images acquired during the period of maximum vegetation activity (end of July – end of August) for the years 2015, 2016 and 2017 were used for the present study.

The acquisition dates and the covered time spans slightly vary among different datasets. Nonetheless, our statistical analyses focused on a steady-state perspective, i.e. considering only the spatial variability of rock glacier status and the associated environmental controls, assuming that discrepancies in time among the various input data are negligible compared to the temporal scale of climate alterations needed to modify rock glacier status. This approach led to an a-priori exclusion of information with a high temporal variability such as weather and climatic data.

4.3. Definition of the response variable

Different terminologies to describe the status of rock glaciers are present in literature. Some inventories distinguish between ice-containing moving (“active”), ice-containing non-moving (“inactive”) and ice-free rock glaciers (e.g. Bollmann et al., 2012), while some others refer only to whether frozen material might be present (“intact”) or not (“relict” or “fossil”) (e.g. Scotti et al., 2013). Different terms are also found to describe “decayed” rock glaciers. Some inventories use the term “fossil” (e.g. Marcer et al., 2017), while others use the term “relict” (e.g. Lilleøren and Etzelmüller, 2011). Barsch (1996) highlighted similarities between these two terms and suggested to use them as synonyms.

Within this study we define the relevant terms for rock glacier characterization as follows: since the work aims to model the likelihood of frozen material being present in rock glaciers, the terms that describe the kinematic behavior of ice-containing rock glaciers “active” and “inactive” were summarized to “intact”. In agreement with most studies, the term “relict” is used to describe a rock glacier that lost most of or entirely its frozen core. The obtained response variable “status of activity” therefore was defined as a binary character.

4.4. Potential predictor variables and data preparation

A crucial step of this study consisted in the selection of potentially suitable variables to estimate the status of a rock glacier. Based on the available literature (cf. Chapter 2), we have identified several predictor variables that may influence the presence of ice within a rock glacier. These variables were subsequently grouped into three thematic fields, namely *topographic controls*, *environmental controls* and *rock glacier appearance* (i.e. the visually perceptible characteristics of a rock glacier). The first three thematic fields summarize assumed drivers for frozen material being present in rock

glaciers. The latter one, *rock glacier appearance*, expresses an effect of ice-presence or -absence within a rock glacier.

All the potential spatial predictor variables described in the following sections and listed in Fig. 3 were resampled to a common grid size of 10 x 10 m and aligned to the same spatial extent and projection. To ensure computational feasibility, rock glacier locations were not represented in a raster mode, rather by a total number of 50,000 sample points randomly distributed within the polygons of the rock glacier inventory.

4.4.1. Topographic controls

The *Geomorphic Protection Index (GPI)* was assumed to best represent the surrounding terrain characteristics of rock glaciers. The index was developed by Yokoyama et al. (2002) and describes the degree of terrain openness. The index calculates the angle between a viewer's standpoint and the highest or lowest visible topographic feature in a given line-of-sight direction. For each standpoint (e.g. grid cell), eight directions within a given radial distance are sampled and averaged. The index summarizes the exposedness of a cell to atmospheric (e.g. incoming solar radiation) or gravitational inputs (e.g. rock fall, snow avalanches).

The *aspect* was considered as a potential predictor because of the known association between slope orientation and incoming solar radiation, which in turn influences permafrost occurrence in mountainous terrain (Harris et al., 2009; Hofierka and Suri, 2002). In order to account for its circular nature and to avoid loss of information due to reclassification, the aspect layer was transformed into “east-exposedness” and “north-exposedness” by performing sine and cosine transformations, respectively (Brenning, 2009).

4.4.2. Environmental controls

In mountainous terrain, the MAAT is directly linked to altitude through the temperature gradient (e.g. a yearly mean of $-0.57^{\circ}\text{C} \cdot 100\text{m}^{-1}$ for the valley slopes of the region Trentino-Alto Adige/South Tyrol, according to Rolland, 2003). This direct relationship between the two parameters requires the selection of one to avoid redundant information within data-driven models (Deluigi et al., 2017). For the present work, *altitude*, i.e. the elevation values stored in the available DTM, was considered as a proxy for elevation-dependent temperature conditions.

A map of the *potential incoming solar radiation (PISR)* was generated in SAGA GIS. The dedicated tool allows to derive spatially-comprehensive radiation values based on topographical as well as positional information (Böhner and Antonić, 2009; Hofierka and Suri, 2002). In order to account only for the snow-free period, PISR was calculated from beginning of June to the end of October and clear sky conditions with an atmospheric transmittance of 70% were assumed as described in Marcer et al. (2017).

Perennial snowfields were selected as potential predictor variable in order to account for potential relationships between perennial snow cover and rock glaciers (i.e. their status). The topographical catchment was computed for each inventory entry and used to identify the perennial snowfields that are present in the contributing area of rock glaciers. To approximate the interactions between rock glaciers and snow deposits, the Euclidean distance from the selected perennial snowfields to the rock glaciers boundary was calculated. The produced spatial variable represents the distance to perennial snowfields in the contributing area of rock glaciers.

As described in Chapter 2, a difference between debris and talus rock glaciers according to their activity status was reported. Therefore, the *sediment origin (talus/till)* of rock glaciers was considered as potential predictor variable.

4.4.3. Rock glacier appearance

A map of *front slope* angles was created by placing various sample points on the front of each rock glacier present in the inventory. The relative slope value of each sampling point has been extracted from an underlying slope map. Then, a median front slope for each rock glacier has been calculated.

The “Terrain Ruggedness Index (TRI)”, developed by Riley et al. (1999), was chosen to represent the undulated morphology that characterizes active rock glaciers, i.e. its *surface relief*. Wahrhaftig and Cox (1959) reported values between 2 and 10 m of height and 2–50 m in distance between ridges on the surface of rock glaciers. In order to approximate this range, the lowest search radius possible has been selected (i.e. 1 grid cell).

The median area of the rock glaciers stored in the South Tyrolean inventory is 96,170 m². In order to capture a general *curvature* of the rock glacier’s surface, plan and profile curvature maps with a moving window radius of 100 m and 500 m were created. As final curvature value, the median of values on the surface excluding the root and the frontal zone of each rock glacier were selected.

The NDVI was assumed to be a suitable predictor in differentiating between intact and relict rock glaciers due to the reported absence of vegetation cover on intact rock glaciers (cf. Chapter 2)

4.5. Construction of the multiple-variable model

Different procedures for the selection of input data for geomorphological predictions have previously been applied and discussed in literature: Brenning (2008) proposed the widely used automatic backward or forward stepwise variable selection that is based on the *Akaike Information Criterion (AIC)*, a measure of goodness-of-fit that penalizes for model complexity (e.g. Goetz et al., 2011; Petschko et al., 2014). Steger et al. (2016) highlighted the drawbacks of these purely number-driven variable selections in a geomorphic modelling context, as they may increase the apparent model performance but produce geomorphologically implausible results. Field et al. (2012, p. 266) support a heuristic approach for variable selection by stating that “*when there is a sound theoretical literature available, then base your model upon what past research tells you*”. We decided to favor the expert-based method over a purely number-focused stepwise variable selection, also to enhance model comparability (i.e. automated selections may produce different variable sets for the different classifiers) and a process orientated model interpretation (cf. Steger et al., 2016).

The conducted explanatory data analysis supported the selection procedure. For each continuous variable, absolute frequencies (i.e. histogram) and relative frequencies (intact vs. relict) were plotted over the respective predictor values. Spineplots were used to visualize relationships among rock glacier status and categorically scaled variables. An interpretation of the plots allowed to explore empirical associations among the response variable and potential predictors as well as their plausibility.

In this study, a variable was selected for being included in the models, when associations that are supported by literature findings (cf. Chapter 2) were supported also by the available data (e.g. conditional frequency plots). The quantitative evaluation of subsequent models is described in Section 4.7.

Many of the considered predictor variables can be considered to be correlated with each other: PISR and NDVI are determined by elevation and aspect (e.g. Brenning and Trombotto, 2006; Deluigi et al., 2017), which again depend on tectonic conditions. Non-independence of variables, i.e. collinearity is an intrinsic effect of most real-world data, because collinear variables are the manifestation of the same – often unmeasurable – processes. The selection of predictor variables should therefore be based on known functional relationships with the response. If the variable selection is reason-

able from a process-perspective, collinearity of variables may not necessarily be problematic (Dormann et al., 2013).

4.6. Classifiers

4.6.1. Generalized linear model (GLM) - Logistic regression

Binomial logistic regression models are popular in many scientific fields to model a binary response (e.g. rock glacier status: intact vs. relict) while allowing to include both, continuous and categorical predictor variables (cf. Section 4.4.). A logistic regression model allows to estimate the likelihood that one class of a binary response (e.g. rock glacier = intact) is present.

In contrast to a linear regression model, logistic regression by its nature does not yield likelihood values outside the interval [0, 1] and therefore meets the formal requirements for probabilistic modelling. In the context of this study, a logistic regression model determines the probability of rock glacier “*intactness*” by a set of predictor variables x , hence $P(Y=1|x)$. The “*odds*”, i.e. the conditional probability of $Y=1$, are transformed to logits with a logit-link function. The *logits*(x) are of a continuous character and therefore a linear model consisting of an intercept α and a predictor variable coefficient β can be fitted upon them.

In a linear regression model, parameters are fitted by a “*least-square method*”, i.e. by minimizing the sum of squared residuals. If the response variable is of a dichotomous character, the least-square method is not applicable. Logistic regression therefore uses a technique called “*maximum likelihood*”. Hereby, values which maximize the probability of obtaining the observed data set are chosen as unknown model parameters (Hosmer and Lemeshow, 2000; Marcer et al., 2017). In R, logistic regression is implemented in the generalized linear model (GLM) function.

Logistic regression models are widely used in the field of earth sciences, for example to model landslide susceptibility (e.g. Schlögel et al., 2018) or the spatial distribution of permafrost (e.g. Sattler et al., 2016).

4.6.2. Support vector machine (SVM)

A SVM is able to identify non-linear relationships between a categorical outcome variable and a set of predictors. The basic principle behind the SVM algorithm is that a hyperplane is fitted into the dataset in order to separate between two or more classes. Vapnik (2006) defines an *optimal hyperplane* as a decision function with maximal margins between the data points to be separated. To build an optimal hyperplane, one does not need the full training data set but only a small subsample of points which are closest to it, i.e. the *support vectors*. Cortes and Vapnik (1995) highlight that the generalization ability is great when the optimal hyperplane can be constructed from a small number of support vectors relative to the size of the training data set. In the presence of non-linearity, a SVM that operates in an arbitrary number of dimensions is built (so called *kernel mapping*). Then, a separating hyperplane can be defined even for highly complicated datasets. One mapping function that was identified to have good generalization properties is the Gaussian radial basis function (RBF) (Ballabio and Sterlacchini, 2012). In the present SVM model, the Gaussian RBF kernel was used owing to the reported good generalization capacity for regional permafrost distribution models (Deluigi et al., 2017). The SVM algorithm is frequently applied in geomorphology to delineate landslide-prone terrain (e.g. Kavzoglu et al., 2014; Peng et al., 2014), while its application in the field of alpine permafrost studies is still rather limited with some exceptions (e.g. Deluigi and Lambiel, 2013)

4.6.3. Random forest (RF)

The RF algorithm was firstly proposed by Breiman (2001) and is based on the concept of decision trees as classifiers. Classifiers

based on single decision trees find a classification rule by a step-wise binary splitting of the data into more homogenous sub-groups. Hereby, the homogeneity of the created subgroups is measured by the Gini impurity index, a metric that measures the probability of an element to be incorrectly classified. A decision tree is assumed to be fully grown when the Gini impurity index does not decrease anymore (Breiman et al., 1984).

The RF algorithm creates a large set of decision trees out of random subsamples of training observations and predictor variables. The remaining data is used to test the fitting performance of each tree. The obtained performance is usually called “out-of-bag” error. The final RF model is an ensemble of decision trees and usually provides a more stable prediction compared to single decision trees as it was shown by Prasad et al. (2006) in the context of tree species distribution modelling. In the geoscientific context, RF was applied by Deluigi et al. (2017) to model spatial permafrost distribution and by Vorpahl et al. (2012) or Catani et al. (2013) to model landslide susceptibility.

4.7. Modelling and model validation

Different methods are available to validate a data-driven spatial prediction model. Still widely used *holdout validation* splits the original data in two subsets, one training set to fit the model and one independent test set on which the previously trained model (e.g. the classification rule) is tested. *K-fold cross validation* (CV) is a more sophisticated approach to evaluate model performance. CV is not reliant on one specific data partition as it splits the full data n times into k random subsamples (also called *folds*) of training and test sets in order to enable a multi-fold model validation (Kohavi, 1995). However, when dealing with georeferenced data, a spatially explicit partitioning of training and test samples is required to avoid overoptimistic performance estimates and to assess the spatial transferability of the model. *Spatial k-fold cross validation* (SCV) splits the initial data set into spatially independent training and test areas via a k -means clustering algorithm (Brenning, 2012). In the present work, the study region was split into 5 sub-regions and SCV was repeated 25 times for each fold (i.e. $25 \times 5 = 125$ tests per model). The discriminatory power of a binary classifier can be measured with the *Area under the Receiving Operating Characteristic* (AUROC). This global performance measure summarizes the portion of correctly classified cases (i.e. true positive rates against the false positive rate for each possible cut-off value). An AUROC of 0.5 relates to a random model while a value of 1 describes a perfect discrimination between the two classes. In the context of the present study, the SCV-based AUROC's express the ability of a model to discriminate between intact and relict rock glaciers of the training samples (fitting performance) or the spatially independent test data (predictive performance).

For both machine learning algorithms, SVM and RF, classification performance depends largely on the careful estimation of model parameters, called hyperparameters. The SVM requires the tuning to two hyperparameters, C (penalty parameter for misclassification) and σ (inverse kernel width). The RF needs the parameters $mtry$ (number of variables randomly sampled as candidates at each split) and $ntree$ (number of trees) to be set. Hyperparameters are ideally identified within an iterative procedure. As suggested by Schratz et al. (2019) a spatial resampling method (i.e. SCV) was applied also during the process of hyperparameter tuning in order to avoid biases due to spatial autocorrelation of sampling points. Within this study, SCV-based hyperparameter tuning was based on a grid search, i.e. an iterative search procedure within a user-defined range of values for C , σ and $mtry$. For the SVM-based model, 256 hyperparameter combinations were tested, ranging from 2^{-12} to 2^3 for both C and σ . A range from 1 to 10 was tested as $mtry$ parameter for the RF-model.

SCV identified values of $C = 0.000976563$ and $\sigma = 0.015625$ as ideal hyperparameters for the SVM. An $mtry$ -value of 1 was determined as hyperparameter for the RF model. The number of grown trees ($ntree$) during model fitting was set to 800.

A permutation-based spatially explicit variable importance was performed to estimate the relevance of each predictor within the models. After randomly permuting (i.e. re-organizing) a predictor variable within a SCV framework, the change in predictive performance was measured via the AUROC metric. The magnitude of decrease in model performance indicates the relative importance of the permuted variable within a multiple variable model. In other words, the observed AUROC decrease allowed insights into the relative importance of each predictor variable. The whole process is repeated for each predictor of the multiple-variable model (Strobl et al., 2007).

4.8. Model application at the grid cell scale

The obtained classification rules (i.e. probability of intact rock glacier for the three models) were transferred to each rock glacier grid cell. For each cell, a continuous number representing the modelled likelihood of the status “intact” was computed. Then, the median value of all the modelled values within a rock glacier was built. An optimal threshold to split the continuous scale of modelled likelihoods into discrete binary labels had to be found. Different methods to identify this optimal cut-point were proposed in literature (see López-Ratón et al., 2014 for an overview). In this study, the point with maximum true positive and true negative rate on the AUROC curve was chosen as described in Riddle and Stratford (1999). The standard deviation allowed insights into the dispersion of spatially predicted probability values for each rock glacier.

The spatially transferred likelihood values (i.e. model output) were randomly sampled ($n=100,000$) and plotted against the respective predictor values. The obtained response plots visualize the trained model as a function of one specified model variable for the respective study area. In addition, they allow a comparison between models produced by different classifiers and their response to a single predictor variable (Lombardo and Mai, 2018; Vorpahl et al., 2012). Finally, the agreement between the three generated model outputs was analyzed by computing the percentage of rock glaciers labelled as either intact or relict by all models or just by two out of three.

5. Results

5.1. Exploratory data analysis and predictor selection

The exploratory data analysis allowed first insights (Fig. 4) into empirical associations between the rock glacier states (intact vs. relict) and potential predictor variables. The conditional density plots depicted that topographically protected slopes (GPI) and higher elevations ($altitude$) were generally associated with a higher portion of intact rock glaciers (Fig. 4). Both variables related to the *slope orientation*, namely north-exposedness and west-exposedness showed no systematic change of intact rock glaciers over the range of slope aspect values. Instead, a higher share of relict rock glaciers with increasing PISR values could be observed. The conditional density plots also show that rock glaciers with *perennial snowfields* in their proximity are more likely to be intact. For the subsequent use within the multiple-variable model, a cutoff value of 300 m of horizontal distance was set, as no effect of perennial snowfields on rock glacier status beyond this value was assumed. The spineplot related to *debris origin* shows that a higher percentage of till-derived rock glaciers is intact compared to their talus-fed counterparts. Steeper *front slope* angles and a higher

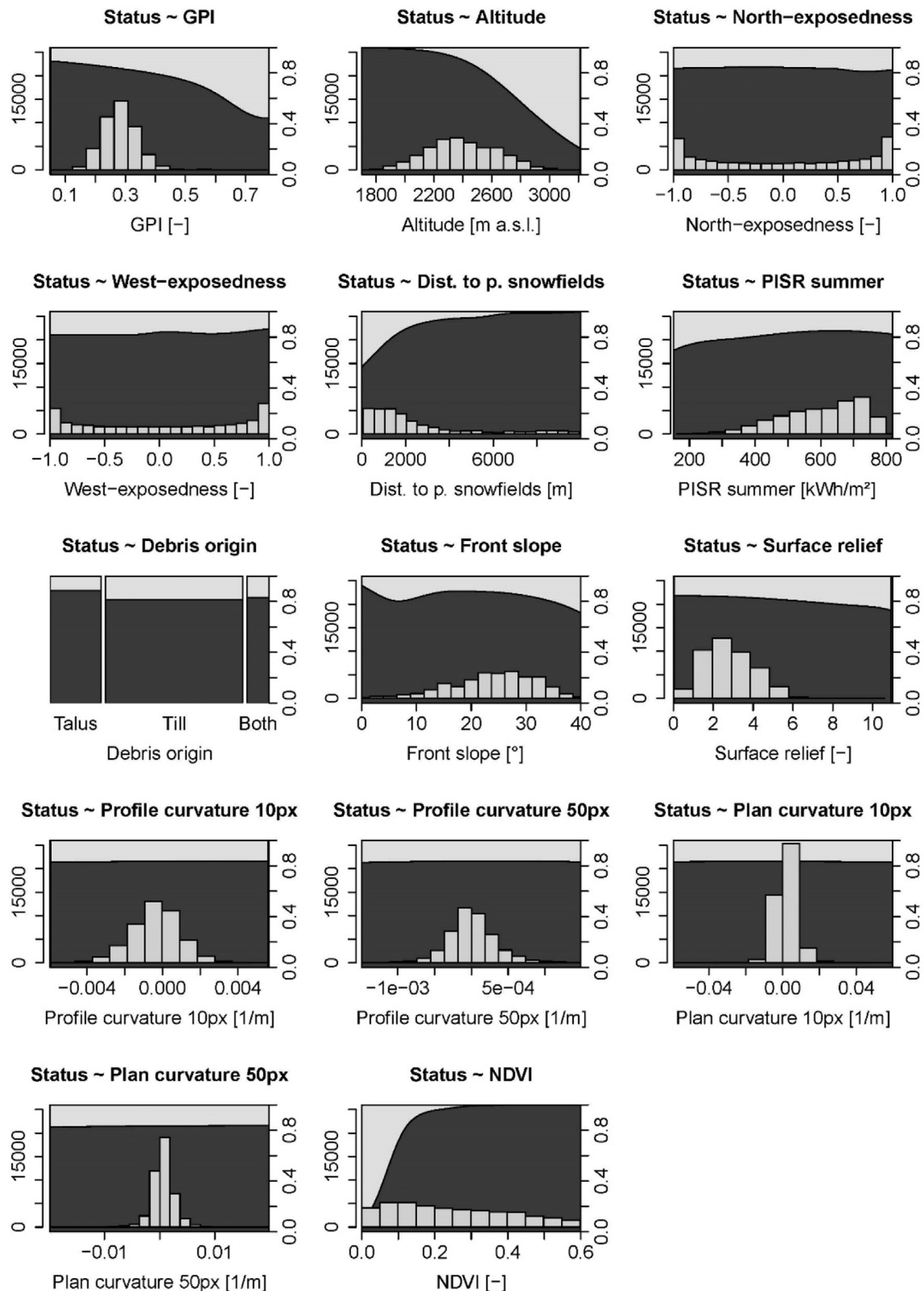


Fig. 4. Absolute frequencies (left y-axis) and conditional distribution (right y-axis) of the potential predictor variables. Shading of the conditional density plots is showing the rock glacier states (intact: light grey; relict: dark grey).

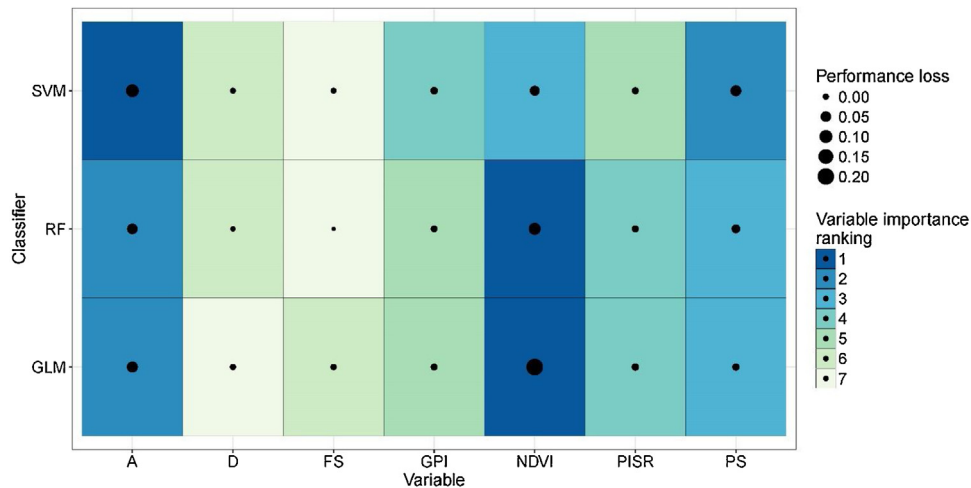
roughness of the surface relief characterize intact rock glaciers, but in all four *curvature* maps the share between intact and relict rock glaciers did not change over the value range. Concerning *NDVI*, the conditional density plot indicates a clear decrease of intact rock glaciers with increasing vegetation activity.

The variables that were selected for the final model were altitude, GPI, PISR, distance to perennial snowfields, front slope angle and debris origin. Their selection is supported by the results of the exploratory data analysis as well as by literature findings. Although *surface relief* has been described as characteristic for active rock

Table 1

Rock glacier status classifications according to the inventory and the three models logistic regression (GLM), support vector machine (SVM) and random forest (RF).

Inventory			GLM			SVM		RF	
Status	n	%	Status	n	%	n	%	n	%
Active	240	13.5	Intact	474	30.7	490	31.7	411	26.6
Inactive	59	3.3	Relict	1070	69.3	1054	68.3	1133	73.4
Relict	1245	70.0	Intact	171	72.8	180	76.6	160	68.1
Unclassified	235	13.2	Relict	64	27.2	55	23.4	75	31.9

**Fig. 5.** Importance ranking of the chosen predictor variables determined by SCV-based variable permutation. Variable abbreviations: A = Altitude, D = Debris origin, FS = Front slope, GPI = Geomorphic protection index, NDVI = Normalized difference vegetation index, PISR = Potential incoming solar radiation, PS = perennial snowfields.

glaciers, this variable was excluded from the analysis due to the uncertainties related to its estimation by the approach deployed here. The variables related to slope aspect and surface curvature were not considered for the multiple variable model because no substantial change of intact rock glaciers over their value range was detected.

5.2. Model performance and validation

Fitting performances for all models achieved a median AUROC higher than 0.95, which can be defined as an outstanding discrimination between intact and relict rock glaciers according to Hosmer and Lemeshow (2000). The RF model reached a near perfect classification of training data, followed by the GLM and the SVM (Table 1). SCV revealed that all models enabled to classify the rock glacier status of spatially independent test data with a very high accuracy (i.e. median AUROC > 0.92). In this context, the RF classifier (AUROC 0.93) achieved a lower performance than the SVM (AUROC 0.95) and the GLM (AUROC 0.96) and revealed a slight tendency to adapt too strongly to the training data set (i.e. to overfit). Differences among median training and test set AUROC's highlighted a high ability of the GLM and SVM based models to generalize (differences < 0.001) while providing indications of a slightly overfitted RF model (difference 0.07). Machine learning techniques (SVM and RF) were associated with higher standard deviations of predictive performance estimates (> 0.029) compared to the more robust GLM (0.002).

Variable permutation based on SCV allowed to rank the importance of each predictor within the multi-variable set of each classifier. A permuted NDVI results in the highest loss of performance in the GLM model, followed by altitude. The two machine learning algorithms in contrast show a less pronounced drop in model performance when disassociating single predictor variables. A permuted NDVI in the RF and SVM results in a loss of about 0.1

and 0.05 AUROC values, respectively. In the SVM, the perennial snowfields result to be even more important than the NDVI. The permuted variables *front slope* and *debris origin* instead, are causing a performance loss close to zero for all the classifiers (Fig. 5).

5.3. Rock glacier status maps and model agreement

The optimal predicted likelihood value (i.e. cut-point) to discriminate between intact and relict rock glaciers were 0.18 for the GLM, 0.23 for the SVM and 0.13 for the RF (cf. Section 4.8.). The SVM classified the most rock glaciers as intact, the RF the least. In the original inventory, 235 i.e. 13.2% of the rock glaciers have an unknown status. All three models classified the majority of the unknown rock glaciers as intact. In agreement with the overall classification, the SVM labeled the highest number of unknown rock glaciers as intact, the RF the lowest. Fig. 6 shows the map application of the classification results. The shown extent is the Zay valley, located in the western part of South Tyrol. While the models generally agree on the status of the shown rock glaciers, they display different degrees of prediction variability.

The vast majority of the rock glaciers in the study area (91.5 %) were allocated to the same classes by all three models (i.e. 60.4 % of all rock glaciers were classified as relict by all models and 31.1 % as intact). Only a minor portion of all the rock glaciers (8.5 %) was classified differently by at least one of the three algorithms. In 4.3 % of the cases two out of three models agreed on the relict status and in 4.2 % of the cases two out of three models agreed on the intact status. Model agreement was higher for already classified rock glaciers than for unclassified ones (i.e. 92.1% versus 87.7 % of rock glaciers that were either classified as intact or relict by all models). Among the unclassified rock glaciers, 21.7 % were equally classified as relict, while for 66 % all models agreed on the relict status. Concerning the already classified rock glaciers instead, 66.3

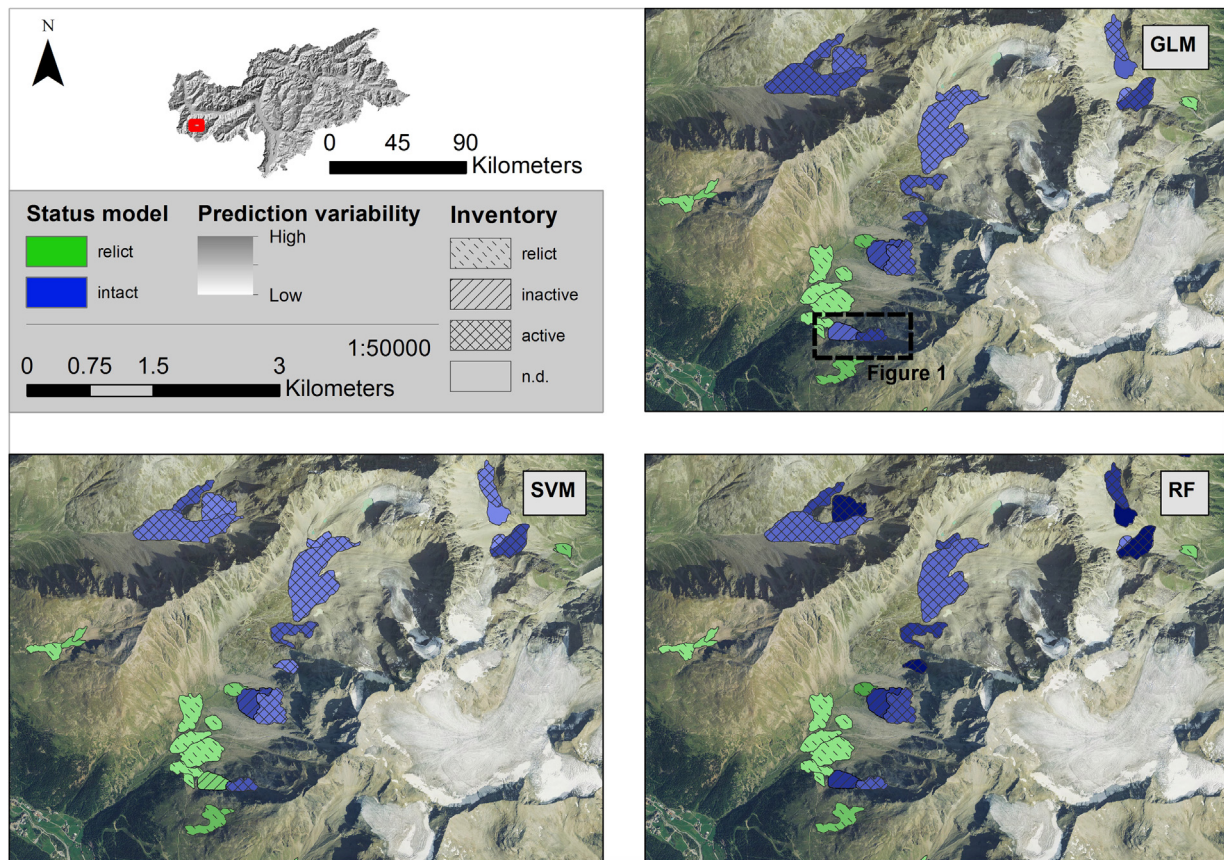


Fig. 6. Modelled rock glacier statuses, variability of predictions and comparison with the inventory for the Zay valley. The rock glaciers highlighted by the dashed box can be viewed in Fig. 1.

% of the rock glaciers were classified as relict by all models, while for 25.8 % all models suggest an intact status (Table 3).

The high level of classification agreement among the three models was particularly evident for the portion of rock glaciers that were already classified in the inventory within altitudes below 2200 m and above 2800 m. Across-model agreement was lower for the unclassified portion of rock glaciers in the inventory, especially below 2200 m. The larger share of model agreement on the intact status for unclassified rock glaciers extends over the whole altitudinal belt (Fig. 7). The density curves of Fig. 7 show that unclassified rock glaciers are more likely to be at higher elevations than their classified counterparts.

The response curves visualized the spatially predicted likelihood (the closer to 1 the more likely intact) along the range of a selected predictor variable (Fig. 8). The altitude plot illustrates how the vast majority of cells located below approximately 2500 m a.s.l. exhibit likelihood values close to 0 for all models, thus a very low modelled probability of a rock glacier being intact. This is in line with empirical frequencies that depict a substantial drop in the number of relict rock glaciers above 2500 m (Fig. 8). The overlay of smoothed response functions allowed a direct visual comparison of the three algorithms, which revealed some differences. The SVM model showed the most distinctive increase of predicted likelihoods per unit of altitude (i.e. steepest curve) confirming the most prominent role of this predictor in the multiple-variable setup, while the response curve related to the RF classifier started to rise at higher altitudes and was constantly below the other two classifiers.

The NDVI-plots showed that above a value of 0.2, model predictions are approaching towards zero-likelihood of intact status. This observation was confirmed by the relative frequency plot of Fig. 4, which showed the highest share of intact rock glaciers

between NDVI-values of 0 and 0.2. Some differences could be noticed between the three applied classifiers. The RF algorithm produced the steepest drop of model outcomes per unit NDVI compared to the GLM and the SVM. The tendency of the SVM to be more optimistic than RF and GLM is reflected here as well as in the number of absolute classifications (see Table 2)

6. Discussion

6.1. Predictor variables and their importance

Altitude or NDVI resulted to be the most important predictor variables in each of the model setups (see Fig. 5). With the altitude as a proxy for the elevation-dependent temperature gradient, one variable expressing a main driver for ice-presence and with the NDVI one variable expressing a visual discernible effect of rock glacier activity resulted to be most important. However, the interplay among predictor variables should be kept in mind, especially with regard to the NDVI, as the Pearson correlation coefficient between NDVI and altitude is -0.74 . Indeed, the extent to which reduced vegetation presence on intact rock glaciers is determined by ice-driven dynamics or simply by a less favorable climate due to higher altitudes was not quantified in this study but should be considered when interpreting the results.

Nevertheless, the NDVI revealed to be a relevant variable to monitor the rock glacier status, also in practical terms. NDVI maps can be produced in a cost-effective way based on freely-available satellite imagery. Due to the high temporal resolution of data provided by recent missions such as the Sentinel-2 constellation, remote sensing based monitoring of the rock glacier status over time may become more effective when taking the NDVI into con-

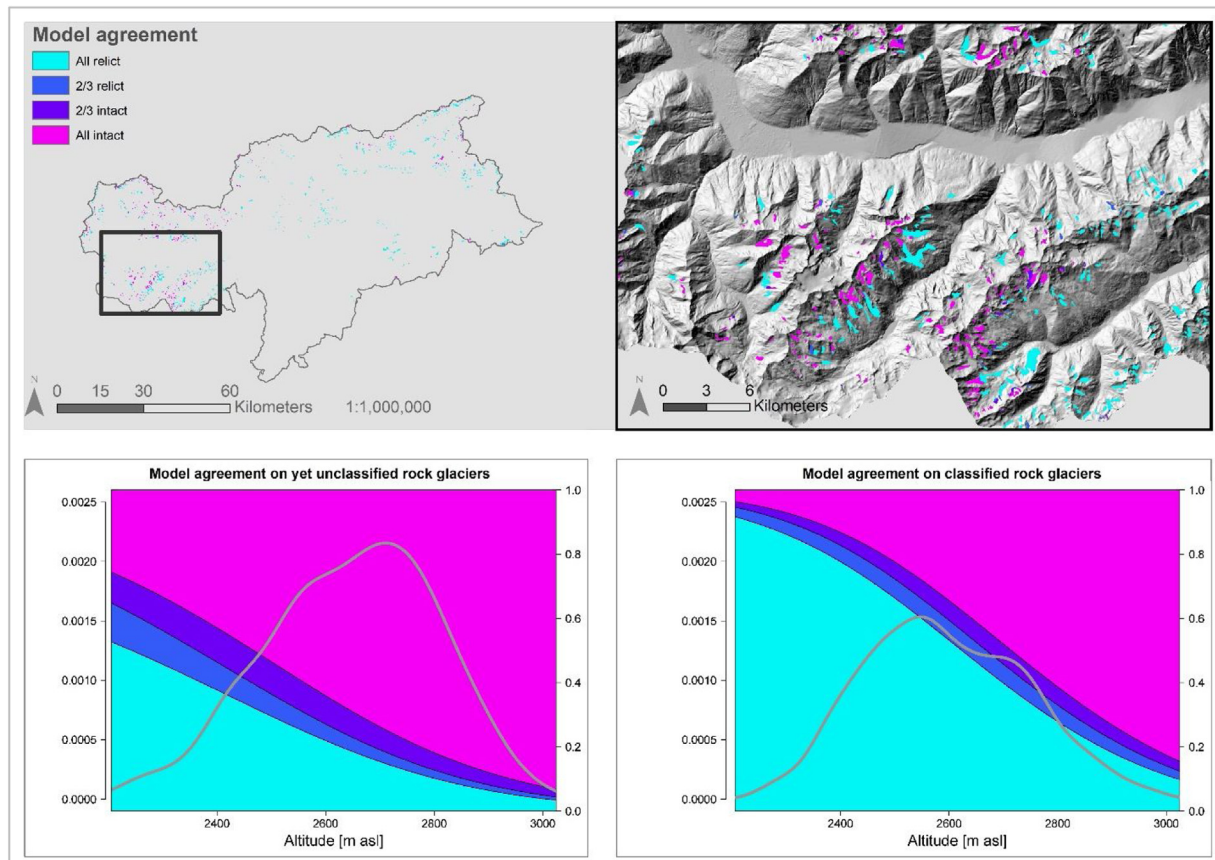


Fig. 7. Model agreement for the entire study region (upper left corner), a detailed extent (upper right corner) and along altitude for unclassified rock glaciers (lower left corner) and classified rock glaciers (lower right corner) present in the inventory for South Tyrol. The grey line in the conditional density plots shows the density distribution of unclassified and classified rock glaciers along altitude.

sideration. However, the regional context has to be considered when using this variable. Outside the European Alps for instance, vegetation growth was found on actively moving rock glaciers, as documented by [Bolch and Gorbunov \(2014\)](#). In single cases, presumably intact rock glaciers can also be covered by a comprehensive vegetation layer in the European Alps as shown recently by [Colucci et al. \(2019\)](#) in the Italian Carnic Alps.

As to topographic data, areas without high-resolution LiDAR-data, DTM's derived from spaceborne synthetic aperture radar (SAR) sensors such as the TanDEM-X data set provided by the German Aerospace Center (DLR) may be a viable alternative.

The generated map of distance to perennial snowfields in the contributing area of rock glaciers achieved a relatively high importance ranking in all three models. This result supports recent findings on how perennial snow and/or avalanche snow might be of high relevance for assessing the rock glacier status (e.g. [Anderson et al., 2018](#)). The available input dataset however, did not allow to reconstruct the genesis of a mapped snowfield, and thus it was not possible to determine whether it was fed by avalanche snow or by another source. Compared to the lower importance-ranking of the GPI, it can be argued that maps of snow deposits on rock glaciers or in their contributing area may display better the process of snow avalanches providing ice and rock debris compared to a proxy variable representing their steep surrounding source areas. The extraction of snow information from remote sensing images might even improve the performance of this variable.

Solar radiation achieved comparably low ranking scores in all three models, causing little performance losses when permuted. Also in this case, unavoidable interrelations between environmental predictor variables are at play. As shown by [Brenning and](#)

[Trombetta \(2006\)](#), PISR may increase with altitude, also affecting the estimated variable importance within a multiple variable modelling setup. Considering the high importance of this variable in previous studies that modelled rock glacier distribution (e.g. [Brenning and Trombetta, 2006](#)), it may be assumed that this variable has a higher suitability to explain their presence rather than their status.

A distinct morphological feature of rock glaciers, the front slope, did not result as important in all created models. Neither a permuted geomorphic protection index caused substantial performance losses, hence the inclusion of morphometric variables to describe rock glaciers or their surroundings brought minor benefit to estimate the rock glacier state for South Tyrol within the multiple variable model setup. Similarly, the debris origin did not turn out to be relevant to be included in the models.

6.2. Model behavior and performances

Altitude and NDVI achieved the highest variable importance ranks in all three model setups. When considering the performance losses, the GLM-based model shows the largest drops caused by altitude and NDVI only, whereas in the machine learning-based models the losses are more equally distributed across all model components. This might be explained by the higher flexibility of the applied machine learning techniques while logistic regression was less suitable to adjust to the “new” data pattern resulting from permutation. The disadvantage of this enhanced flexibility of the machine learning algorithms is their tendency to adapt too strongly to the training data while failing to predict independent test sets. A confrontation of model performance estimates based

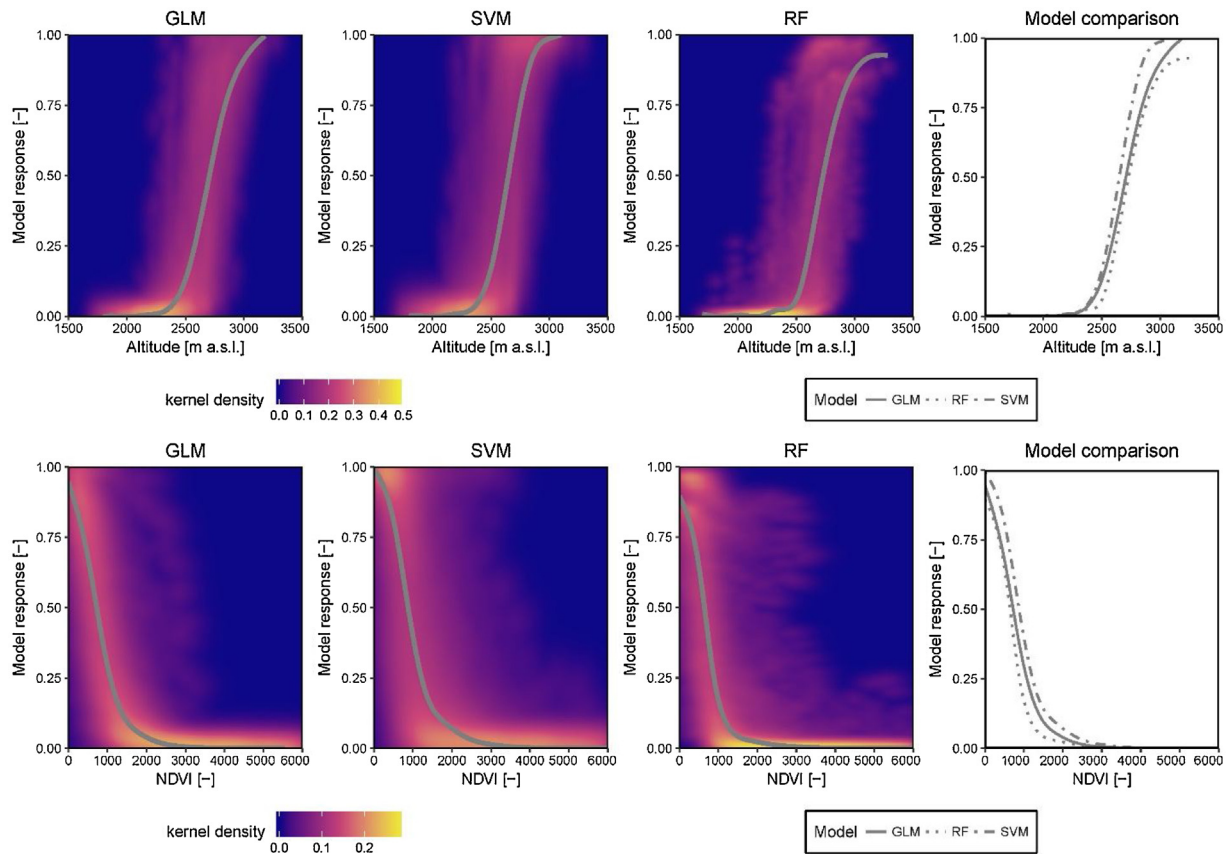


Fig. 8. Model response for the GLM (1st column), SVM (2nd column) and RF (3rd column) over predictor variables altitude (1st row) and NDVI (2nd row). A direct comparison of model responses is given within the 4th column. A value closer to 1 indicates a higher likelihood of an intact status.

Table 2

Model agreement between unclassified ($n = 235$) and classified ($n = 1544$) entries of the rock glacier inventory.

	Unclassified		Classified		Total	
	n	%	n	%	n	%
All relict	51	21.7	1024	66.3	1075	60.4
2/3 relict	12	5.1	64	4.1	76	4.3
2/3 intact	17	7.2	57	3.7	74	4.2
All intact	155	66.0	399	25.8	554	31.1

Table 3

Median fitting and predictive performance of the classifiers logistic regression (GLM), support vector machine (SVM) and random forest (RF) and associated standard deviation. The overfitting indicator was calculated by subtracting the predictive from the fitting performance.

Performance	Classifier					
	GLM		SVM		RF	
	AUROC	st. dev	AUROC	st. dev	AUROC	st. dev
Training set (fitting performance)	0.963	0.0002	0.954	0.0091	0.995	0.0016
Test set (predictive performance)	0.963	0.0019	0.954	0.0290	0.927	0.0629
Overfitting indicator	0.00	–	0.00	–	0.07	–

on training sets and test sets (fitting performance vs. predictive performance) can give valuable insights into the degree of model overfitting (cf. Table 1) (Steger and Kofler, 2019). Considering the fitting performance of the applied classifiers, RF reaches a nearly perfect separation of intact and relict rock glaciers of the training set (AUROC-score of nearly 1). The simultaneous presence of lowest predictive performance showed that the RF classifier adapted itself too closely to the training data set, also indicating its reduced capacity to correctly classify out-of-model data. GLM and SVM in contrast

showed a good ability to apply fitted models to an independent validation dataset, which led to the conclusion that they are better suited to generalize a model over a regional scale. In particular, the good performance of the GLM may be attributed to the monotonically increasing relationships between rock glaciers status and its most important predictor variables (see also Fig. 4). Flexible, non-parametric learning algorithms seemed not to provide a substantial advantage for our purpose, which rather required a good ability to generalize over a large area. This finding agrees with other studies

highlighting that when generalization is needed, simple statistical methods can perform equally well as complex machine learning models (Brenning, 2005; Vorpahl et al., 2012).

6.3. Interpretation of the generated maps and model limitations

The generated maps shown in Fig. 6 are the spatial expression of the modelled relationships that were calibrated on the basis of the rock glacier inventory of South Tyrol and validated with spatially-independent test samples. However, the inventory was mainly compiled by manually mapping rock glaciers, visually interpreting morphological features that indicate the respective status. Only for a small portion of the rock glaciers stored in the inventory, permafrost presence could be verified directly by detailed in-situ investigations like borehole surveys or spring-temperature measurements (Krainer et al., 2012, 2015a, 2015b). Therefore, it is important to note that the trained models relied strongly on the expert-based assignment of the rock glacier status, which was based on the evaluation of Bollmann et al. (2012).

To divide the continuous model output (i.e. likelihood between 0 and 1) into two classes, an optimal cut-point had to be found. This allowed a comparison with the original inventory as well as a labelling of the yet unclassified rock glaciers into either intact or relict. However, it can be discussed if a discrete map-representation of a rock glacier's status is more appropriate than a continuous one. From a process-oriented perspective, the present-day rock glacier status can be considered as a snapshot in a continuum of climatic variability over time. Hence, a continuous depiction of the status might seem to be more appropriate than a categorical classification. Considering the completion of a rock glacier inventory as one objective of this study, a discrete map-representation including the variability of likelihood-values within one rock glacier was preferred.

An interpretation of the point-density plots and the relative response curves (Fig. 8) allowed insight into the model behavior. The highest increase of modelled likelihood values per unit elevation of the SVM model provided evidence of its high sensitivity to altitude. Such a different sensitivity to this important predictor might be the main reason why GLM and SVM predicted a higher portion of rock glaciers intact, compared to the RF classifier. The RF-model resulted in the highest concentration of values around zero likelihood of activity, both for altitude and NDVI. This suggests that the RF algorithm, due to its concept based on decision trees, tends to avoid immediate probabilities between 0 and 1. In the smoothed response curves, this trend is expressed again by a lower increase of likelihood per unit altitude and a higher drop of likelihood per unit NDVI.

7. Conclusion

The present study highlighted the potential of statistical and machine learning algorithms to model and analyze the status of rock glaciers and its relationship with environmental control variables. Remarkably, highly flexible machine learning classifiers did not perform better than a well-established statistically-based model. It can be concluded that for classification tasks that require a high generalization ability, statistical classifiers might be the preferred choice. It was shown that a successful classification furthermore requires an in-depth preliminary exploratory data examination as well as a spatially meaningful model validation. If machine-learning is applied for classification, hyperparameters should be tuned in a spatial context. Future studies may consider a general additive model (GAM) as a trade-off between linear statistical classifiers and machine learning techniques due to its ability to model nonlinear relationships (Brenning, 2008; Hastie, 2006;

Hastie and Tibshirani, 1990). The consideration of a nested data structure with a mixed-effects model (e.g. with the R-packages “glmm” by Knudson et al., 2018 or “mgcv” by Wood, 2019) may further improve modelling results. Its application in this study was excluded in favor of a comparable presentation of results between the statistical and machine learning models.

The practical applicability of the proposed method was demonstrated by assigning a status to 235 yet unclassified rock glaciers stored in the regional inventory for South Tyrol, Italy. The presented model results are promising in terms of model performance (testing AUROC > 0.9) and geomorphologic plausibility. The three generated model setups generally agreed among each other by classifying more than 80% of the rock glaciers stored in the South Tyrolean inventory as either intact or relict. Only for a small portion of rock glaciers, one model out of three suggested a different label.

Despite the satisfactory model performance and agreement among applied classifiers, one should be aware that the presented method does not prove the evidence for the status of a given rock glacier. Indeed, the proposed method is affected by several limitations. First, rock glacier status was treated as static in this work, and this is a strong simplification of a process that is characterized by temporal dynamics. The limitation hereby arises from the static character of GIS and maps in general: the inventory has to be considered as a snap-shot in time which was reproduced by the applied classifiers. The same shortcoming is to be considered with regard to the applied predictor variables.

Secondly, the method relies on the quality of the rock glacier inventory, particularly the validity of the expert-based assignment of the rock glacier status. A careful assessment of the rock glacier status by the operator in charge to generate it is crucial. Before applying the presented workflow, one should be aware of eventual systematic biases included in the inventory. Such errors could be introduced for example if the task of mapping and interpreting rock glaciers is assigned to more than one operator (Brardinoni et al., 2019; Schmid et al., 2015).

On the other hand, our method can be reproduced in data-poor regions as NDVI and elevation – the most important predictor variables for the rock glacier status – can be generated based on globally and often freely available satellite data. In this context, the development of methods that are less reliant on the availability of an inventory should be subject of further studies.

Acknowledgements

This research was realized within the PhD-work of Christian Kofler, which is kindly funded through a scholarship provided by Stiftung Südtiroler Sparkasse/Fondazione Cassa di Risparmio di Bolzano and Eurac Research. Mattia Rossi is acknowledged for processing Sentinel-2 images. We thank three anonymous reviewers for their constructive comments on the manuscript.

References

- Anderson, R.S., Anderson, L.S., Armstrong, W.H., Rossi, M.W., Crump, S.E., 2018. Glaciation of alpine valleys: the glacier – debris-covered glacier – rock glacier continuum. *Geomorphology* 311, 127–142. <http://dx.doi.org/10.1016/j.geomorph.2018.03.015>.
- Arenson, L., (Doctoral Thesis) 2002. *Unstable Alpine Permafrost: A Potentially Important Natural Hazard*. ETH Zürich, Zürich.
- Azócar, G.F., Brenning, A., Bodin, X., 2017. Permafrost distribution modelling in the semi-arid Chilean Andes. *Cryosphere* 11, 877–890. <http://dx.doi.org/10.5194/tc-11-877-2017>.
- Ballabio, C., Sterlacchini, S., 2012. Support Vector Machines for Landslide Susceptibility Mapping: The Staffora River Basin Case Study, Italy. *Math. Geosci.* 44, 47–70. <http://dx.doi.org/10.1007/s11004-011-9379-9>.
- Barboux, C., Delaloye, R., Lambiel, C., 2014. Inventorying slope movements in an Alpine environment using DInSAR. *Earth Surf. Process. Landf.* 39, 2087–2099. <http://dx.doi.org/10.1002/esp.3603>.

- Barsch, D., 1996. *Rockglaciers - Indicators for the Present and Former Geocology in High Mountain Environments*. Springer Series in Physical Environment. Springer, Berlin Heidelberg, Heidelberg.
- Berthling, I., 2011. Beyond confusion: rock glaciers as cryo-conditioned landforms. *Geomorphology* 131, 98–106, <http://dx.doi.org/10.1016/j.geomorph.2011.05.002>.
- Berthling, I., Etzelmüller, B., 2011. The concept of cryo-conditioning in landscape evolution. *Quat. Res.* 75, 378–384, <http://dx.doi.org/10.1016/j.yqres.2010.12.011>.
- Bertone, A., Zucca, F., Marin, C., Notarnicola, C., Cuozzo, G., Krainer, K., Mair, V., Riccardi, P., Callegari, M., Seppi, R., 2019. An unsupervised method to detect rock glacier activity by using Sentinel-1 SAR interferometric coherence: a regional-scale study in the northern European Alps. *Remote Sens.* 11, 1711, <http://dx.doi.org/10.3390/rs11141711>.
- Blöthe, J.H., Rosenwinkel, S., Höser, T., Korup, O., 2018. Rock-glacier dams in high Asia. *Earth Surf. Process. Landf.*, <http://dx.doi.org/10.1002/esp.4532>.
- Boeckli, L., Brenning, A., Gruber, S., Noetzi, J., 2012. A statistical approach to modelling permafrost distribution in the European Alps or similar mountain ranges. *Cryosphere* 6, 125–140, <http://dx.doi.org/10.5194/tc-6-125-2012>.
- Böhner, J., Antoni, O., 2009. Land-surface parameters specific to topo-climatology. In: Hengl, T., Reuter, H.I. (Eds.), *Developments in Soil Science, Geomorphometry*. Elsevier, pp. 195–226, [http://dx.doi.org/10.1016/S0166-2481\(08\)00008-1](http://dx.doi.org/10.1016/S0166-2481(08)00008-1).
- Bolch, T., Gorbunov, A.P., 2014. Characteristics and origin of rock glaciers in Northern Tien Shan (Kazakhstan/Kyrgyzstan): rock glaciers in Northern Tien Shan. *Permafrost. Periglac. Process.* 25, 320–332, <http://dx.doi.org/10.1002/ppp.1825>.
- Bollmann, E., Rieg, L., Spross, M., Sailer, R., Bucher, K., Maukisch, M., Monreal, M., Zischg, A., Mair, V., Lang, K., Stötter, J., 2012. *Blockgletscherkataster in Südtirol - Erstellung und Analyse*. In: Stötter, J., Sailer, R. (Eds.), *Permafrost in Südtirol, Innsbrucker Geographische Studien*. University of Innsbruck, Innsbruck, Austria, pp. 147–171.
- Brardinoni, F., Scotti, R., Sailer, R., Mair, V., 2019. Evaluating sources of uncertainty and variability in rock glacier inventories. *Earth Surf. Process. Landforms*, <http://dx.doi.org/10.1002/esp.4674>, esp.4674.
- Breiman, L., 2001. *Random forests*. *Mach. Learn.* 45, 5–32.
- Breiman, L., Friedman, J., Stone, C.J., Olshen, R.A., 1984. *Classification and Regression Trees*, 1st ed. Chapman and Hall/CRC Press.
- Brenning, A., 2012. Spatial cross-validation and bootstrap for the assessment of prediction rules in remote sensing: the R package *sperrorst*. In: Presented at the IGARSS 2012 - IEEE International Geoscience and Remote Sensing Symposium, IEEE, Munich, Germany, pp. 5372–5375, <http://dx.doi.org/10.1109/IGARSS.2012.6352393>.
- Brenning, A., 2009. Benchmarking classifiers to optimally integrate terrain analysis and multispectral remote sensing in automatic rock glacier detection. *Remote Sens. Environ.* 113, 239–247, <http://dx.doi.org/10.1016/j.rse.2008.09.005>.
- Brenning, A., 2008. Statistical geocomputing combining R and SAGA: the example of landslide susceptibility analysis with generalized additive models. *Hamburger Beiträge zur Physischen Geographie und Landschaftsökologie* 19, 23–32.
- Brenning, A., 2005. Spatial prediction models for landslide hazards: review, comparison and evaluation. *Nat. Hazards Earth Syst. Sci.* 5, 853–862, <http://dx.doi.org/10.5194/nhess-5-853-2005>.
- Brenning, A., 2003. La importancia de los glaciares de escombros en los sistemas geomorfológicos e hidrológicos de la cordillera de Santiago: fundamentos y primeros resultados. *Rev. Geogr. Norte Gd.* 30, 7–22.
- Brenning, A., Grasser, M., Friend, D.A., 2007. Statistical estimation and generalized additive modeling of rock glacier distribution in the San Juan Mountains, Colorado, United States. *J. Geophys. Res.* 112, <http://dx.doi.org/10.1029/2006JF000528>.
- Brenning, A., Trombotto, D., 2006. Logistic regression modeling of rock glacier and glacier distribution: topographic and climatic controls in the semi-arid Andes. *Geomorphology* 81, 141–154, <http://dx.doi.org/10.1016/j.geomorph.2006.04.003>.
- Catani, F., Lagomarsino, D., Segoni, S., Tofani, V., 2013. Landslide susceptibility estimation by random forests technique: sensitivity and scaling issues. *Nat. Hazards Earth Syst. Sci.* 13, 2815–2831, <http://dx.doi.org/10.5194/nhess-13-2815-2013>.
- Colucci, R.R., Boccali, C., Žebre, M., Guglielmin, M., 2016. Rock glaciers, protalus ramparts and pronival ramparts in the south-eastern Alps. *Geomorphology* 269, 112–121, <http://dx.doi.org/10.1016/j.geomorph.2016.06.039>.
- Colucci, R.R., Forte, E., Žebre, M., Maset, E., Zanettini, C., Guglielmin, M., 2019. Is that a relict rock glacier? *Geomorphology* 330, 177–189, <http://dx.doi.org/10.1016/j.geomorph.2019.02.002>.
- Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., Böhner, J., 2015. System for automated geoscientific analyses (SAGA) v. 2.1.4. *Geosci. Model. Dev.* 8, <http://dx.doi.org/10.5194/gmd-8-1991-2015>.
- Cortes, C., Vapnik, V., 1995. Support-vector networks. *Mach. Learn.* 20, 273–297, <http://dx.doi.org/10.1007/BF00994018>.
- Cossart, E., Fort, M., Bourles, D., Carcaillet, J., Perrier, R., Siame, L., Braucher, R., 2010. Climatic significance of glacier retreat and rockglaciers re-assessed in the light of cosmogenic dating and weathering rind thickness in Clarée valley (Briançonnais, French Alps). *CATENA* 80, 204–219, <http://dx.doi.org/10.1016/j.catena.2009.11.007>.
- Cremonese, E., Gruber, S., Phillips, M., Pogliotti, P., Boeckli, L., Noetzi, J., Suter, C., Bodin, X., Crepaz, A., Kellerer-Pirklbauer, A., Lang, K., Letey, S., Mair, V., Morra di Cella, U., Ravel, L., Scapozza, C., Seppi, R., Zischg, A., 2011. Brief Communication: an inventory of permafrost evidence for the European Alps. *Cryosphere* 5, 651–657, <http://dx.doi.org/10.5194/tc-5-651-2011>.
- Delaloye, R., Strozzi, T., Lambiel, C., Perruchoud, E., Raetz, H., 2008. Landslide-like development of rockglaciers detected with ERS-1/2 SAR interferometry. In: Presented at the 8th International Conference on Permafrost, Swiss Federal Institute of Technology (ETHZ), University of Zürich, Zürich, Switzerland, pp. 26–30.
- Della Chiesa, S., Bertoldi, G., Niedrist, G., Obojes, N., Endrizzi, S., Albertson, J.D., Wohlfahrt, G., Hörtnagl, L., Tappeiner, U., 2014. Modelling changes in grassland hydrological cycling along an elevational gradient in the Alps. *Ecohydrology* 7, 1453–1473, <http://dx.doi.org/10.1002/eco.1471>.
- Deluigi, N., Lambiel, C., 2013. PERMAL: a machine learning approach for alpine permafrost distribution modeling. In: Graf, C. (Ed.), *Presented at the Jahrestagung der schweizerischen geomorphologischen Gesellschaft*. University of Basel, Basel, Switzerland.
- Deluigi, N., Lambiel, C., Kanevski, M., 2017. Data-driven mapping of the potential mountain permafrost distribution. *Sci. Total Environ.* 590–591, 370–380, <http://dx.doi.org/10.1016/j.scitotenv.2017.02.041>.
- Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., McClean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., Lautenbach, S., 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 36, 27–46, <http://dx.doi.org/10.1111/j.1600-0587.2012.07348.x>.
- ESRI, 2016. *ArcGIS Desktop*. Environmental Systems Research Institute, Redlands, CA.
- Field, A., Miles, J., Field, Z., 2012. *Discovering Statistics Using R*. SAGE Publications, Los Angeles, London, New Delhi, Singapore, Washington DC.
- Frauenfelder, R., Kääb, A., 2000. Towards a palaeoclimatic model of rock-glacier formation in the Swiss Alps. *Ann. Glaciol.* 31.
- Giardino, J.R., Vitek, J.D., 1988. The significance of rock glaciers in the glacial-periglacial landscape continuum. *J. Quat. Sci.* 3, 97–103, <http://dx.doi.org/10.1002/jqs.3390030111>.
- Goetz, J.N., Guthrie, R.H., Brenning, A., 2011. Integrating physical and empirical landslide susceptibility models using generalized additive models. *Geomorphology* 129, 376–386, <http://dx.doi.org/10.1016/j.geomorph.2011.03.001>.
- Gruber, S., 2012. Derivation and analysis of a high-resolution estimate of global permafrost zonation. *Cryosphere* 6, 221–233, <http://dx.doi.org/10.5194/tc-6-221-2012>.
- Haeblerli, W., 1985. *Creep of Mountain Permafrost: Internal Structure and Flow of Alpine Rock Glaciers*. Mitteilungen Der Versuchsanstalt Für Wasserbau. Hydrologie und Glaziologie an der Eidgenössischen Technischen Hochschule Zürich, Zürich.
- Haeblerli, W., Hallet, B., Arenson, L., Elconin, R., Humlum, O., Kääb, A., Kaufmann, V., Ladanyi, B., Matsuoka, N., Springman, S., Mühl, D.V., 2006. Permafrost creep and rock glacier dynamics. *Permafrost. Periglac. Process.* 17, 189–214, <http://dx.doi.org/10.1002/ppp.561>.
- Harris, C., Arenson, L.U., Christiansen, H.H., Etzelmüller, B., Frauenfelder, R., Gruber, S., Haeblerli, W., Hauck, C., Hölzle, M., Humlum, O., Isaksen, K., Kääb, A., Kern-Lütschg, M.A., Lehning, M., Matsuoka, N., Murton, J.B., Nötzli, J., Phillips, M., Ross, N., Seppälä, M., Springman, S.M., Vonder Mühl, D., 2009. Permafrost and climate in Europe: monitoring and modelling thermal, geomorphological and geotechnical responses. *Earth. Rev.* 92, 117–171, <http://dx.doi.org/10.1016/j.earscirev.2008.12.002>.
- Hastie, T., 2006. *The Gam Package*.
- Hastie, T., Tibshirani, R., 1990. *Generalized additive models*. Chapman and Hall/CRC Press.
- Hijmans, R.J., 2017. *Raster: Geographic Data Analysis and Modeling*.
- Hoelzle, M., 1992. Permafrost occurrence from BTS measurements and climatic parameters in the eastern Swiss Alps. *Permafrost. Periglac. Process.* 3, 143–147, <http://dx.doi.org/10.1002/ppp.3430030212>.
- Hoelzle, M., Haeblerli, W., 1995. Simulating the effects of mean annual air-temperature changes on permafrost distribution and glacier size: an example from the Upper Engadin, Swiss Alps. *Ann. Glaciol.* 21, 399–405, <http://dx.doi.org/10.3189/S026030550001613X>.
- Hofierka, J., Suri, M., 2002. *The solar radiation model for Open source GIS: implementation and applications*. In: Presented at the Open Source GIS - GRASS Users Conference 2002, Trento, Italy, pp. 51–70.
- Hosmer, D.W., Lemeshow, S., 2000. *Applied logistic regression*. In: *Wiley Series in Probability and Statistics*, 2nd ed. John Wiley & Sons, Inc.
- Humlum, O., Christiansen, H.H., Juliussen, H., 2007. Avalanche-derived rock glaciers in Svalbard. *Permafrost. Periglac. Process.* 18, 75–88, <http://dx.doi.org/10.1002/ppp.580>.
- Ikeda, A., Matsuoka, N., 2006. Pebbly versus bouldery rock glaciers: morphology, structure and processes. *Geomorphology* 73, 279–296, <http://dx.doi.org/10.1016/j.geomorph.2005.07.015>.
- Johnson, B.G., Thackray, G.D., Van Kirk, R., 2007. The effect of topography, latitude, and lithology on rock glacier distribution in the Lemhi Range, Central Idaho, USA. *Geomorphology* 91, 38–50, <http://dx.doi.org/10.1016/j.geomorph.2007.01.023>.
- Johnson, P.G., 1980. Glacier-Rock Glacier Transition in the Southwest Yukon Territory, Canada. *Arct. Alp. Res.* 12, 195, <http://dx.doi.org/10.2307/1550516>.
- Jones, D.B., Harrison, S., Anderson, K., Betts, R.A., 2018a. Mountain rock glaciers contain globally significant water stores. *Sci. Rep.* 8, <http://dx.doi.org/10.1038/s41598-018-21244-w>.
- Jones, D.B., Harrison, S., Anderson, K., Selley, H.L., Wood, J.L., Betts, R.A., 2018b. The distribution and hydrological significance of rock glaciers in the Nepalese Himalaya. *Glob. Planet. Change* 160, 123–142, <http://dx.doi.org/10.1016/j.gloplacha.2017.11.005>.

- Jones, D.B., Harrison, S., Anderson, K., Whalley, W.B., 2019. Rock glaciers and mountain hydrology: a review. *Earth. Rev.* 193, 66–90, <http://dx.doi.org/10.1016/j.earscirev.2019.04.001>.
- Karatzoglou, A., Smola, A., Hornik, K., Zeileis, A., 2004. kernlab - An S4 Package for Kernel Methods in R. *J. Stat. Softw.* 11, <http://dx.doi.org/10.18637/jss.v011.i09>.
- Kavzoglu, T., Sahin, E.K., Colkesen, I., 2014. Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression. *Landslides* 11, 425–439, <http://dx.doi.org/10.1007/s10346-013-0391-7>.
- Kellerer-Pirklbauer, A., 2007. Lithology and the distribution of rock glaciers: Niedere Tauern Range, Styria, Austria. *Zeitschrift für Geomorphologie* 51, 17–38, <http://dx.doi.org/10.1127/0372-8854/2007/0051S2-0017>.
- Knight, J., Harrison, S., Jones, D.B., 2019. Rock glaciers and the geomorphological evolution of deglaciating mountains. *Geomorphology* 324, 14–24, <http://dx.doi.org/10.1016/j.geomorph.2018.09.020>.
- Knudson, C., Geyer, C.J., Benson, S., 2018. Package “glm”.
Kohavi, R., 1995. A study of Cross-validation and bootstrap for accuracy estimation and model selection. In: *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*. Presented at the International Joint Conferences on Artificial Intelligence Organization, International Joint Conferences on Artificial Intelligence (IJCAI), Montreal, Canada, p. 7.
- Krainer, K., Ausserer, P., Bressan, D., Lang, K., Mair, V., Mussner, L., Nickus, U., Schmidt, V., Schiestl, E.-M., Tessadri, R., Thies, H., Tonidandel, D., 2015a. Aufbau und Dynamik ausgewählter Blockgletscher in Nord- und Südtirol. *GeoAlp* 12, 75–134.
- Krainer, K., Bressan, D., Dietre, B., Haas, J.N., Hajdas, I., Lang, K., Mair, V., Nickus, U., Reidl, D., Thies, H., Tonidandel, D., 2015b. A 10,300-year-old permafrost core from the active rock glacier Lazaun, southern Ötztal Alps (South Tyrol, Northern Italy). *Quat. Res.* 83, 324–335, <http://dx.doi.org/10.1016/j.yqres.2014.12.005>.
- Krainer, K., Mussner, L., Behm, M., Hausmann, H., 2012. Multi-disciplinary investigation of an active rock glacier in the Sella Group (Dolomites, Northern Italy). *Austrian J. Earth Sci.*, 105.
- Kummert, M., Delaloye, R., 2018. Regional-scale inventory of periglacial moving landforms connected to the torrential network system. *Geogr. Helv.* 73, 357–371, <http://dx.doi.org/10.5194/gh-73-357-2018>.
- Kummert, M., Delaloye, R., Braillard, L., 2017. Erosion and sediment transfer processes at the front of rapidly moving rock glaciers: systematic observations with automatic cameras in the Western Swiss Alps. *Permafr. Periglac. Process.*, <http://dx.doi.org/10.1002/ppp.1960>.
- Liaw, A., Wiener, M., 2002. Classification and regression by random forest. *R news* 2, 18–22.
- Lilleøren, K.S., Etzel Müller, B., 2011. A regional inventory of rock glaciers and ice-cored moraines in Norway. *Geogr. Ann. Ser. A Phys. Geogr.* 93, 175–191, <http://dx.doi.org/10.1111/j.1468-0459.2011.00430.x>.
- Lombardo, L., Mai, P.M., 2018. Presenting logistic regression-based landslide susceptibility results. *Eng. Geol.* 244, 14–24, <http://dx.doi.org/10.1016/j.enggeo.2018.07.019>.
- López-Ratón, M., Rodríguez-Álvarez, M.X., Suárez, C.C., Sampedro, F.G., 2014. Optimal cutpoints: an R Package for selecting optimal cutpoints in diagnostic tests. *J. Stat. Softw.* 61, <http://dx.doi.org/10.18637/jss.v061.i08>.
- Marazzi, S., 2005. *Atlante Orografico Delle Alpi*. SOIUSA, Priuli & Verlucca, Editori, Scarmagno, Italy.
- Marcet, M., Bodin, X., Brenning, A., Schoeneich, P., Charvet, R., Gottardi, F., 2017. Permafrost Favorability Index: spatial modeling in the French Alps Using a Rock Glacier Inventory. *Front. Earth Sci.* 5, <http://dx.doi.org/10.3389/feart.2017.00105>.
- Martin, H.E., Whalley, W.B., 1987. Rock glaciers: part 1: rock glacier morphology: classification and distribution. *Prog. Phys. Geogr.* 11, 260–282.
- Matsuoka, N., Ikeda, A., 2001. Geological Control on the Distribution and Characteristics of Talus-Derived Rock Glaciers, Vol. 27. Annual Report of the Institute of Geoscience, University of Tsukuba, pp. 11–16.
- Mergili, M., Kerschner, H., 2015. Gridded precipitation mapping in mountainous terrain combining GRASS and R. *Norsk Geografisk Tidsskrift – Norwegian. J. Geogr.* 69, 2–17, <http://dx.doi.org/10.1080/00291951.2014.992807>.
- Monnier, S., Camerlynck, C., Rejiba, F., Kinnard, C., Feuillet, T., Dhemaied, A., 2011. Structure and genesis of the Thabor rock glacier (Northern French Alps) determined from morphological and ground-penetrating radar surveys. *Geomorphology* 134, 269–279, <http://dx.doi.org/10.1016/j.geomorph.2011.07.004>.
- Monnier, S., Kinnard, C., 2015. Reconsidering the glacier to rock glacier transformation problem: new insights from the central Andes of Chile. *Geomorphology* 238, 47–55, <http://dx.doi.org/10.1016/j.geomorph.2015.02.025>.
- Müller, J., Gärtner-Roer, I., Kenner, R., Thee, P., Morche, D., 2014. Sediment storage and transfer on a periglacial mountain slope (Corvatsch, Switzerland). *Geomorphology* 218, 35–44, <http://dx.doi.org/10.1016/j.geomorph.2013.12.002>.
- Peng, L., Niu, R., Huang, B., Wu, X., Zhao, Y., Ye, R., 2014. Landslide susceptibility mapping based on rough set theory and support vector machines: a case of the Three Gorges area, China. *Geomorphology* 204, 287–301, <http://dx.doi.org/10.1016/j.geomorph.2013.08.013>.
- Petschko, H., Brenning, A., Bell, R., Goetz, J., Glade, T., 2014. Assessing the quality of landslide susceptibility maps – case study Lower Austria. *Nat. Hazards Earth Syst. Sci.* 14, 95–118, <http://dx.doi.org/10.5194/nhess-14-95-2014>.
- Potter, N., 1972. Ice-cored rock glacier, Galena Creek, Northern Absaroka Mountains. Wyoming. *Geol. Soc. Am. Bull.* 83, 3025, [http://dx.doi.org/10.1130/0016-7606\(1972\)83\[3025:IRGGCN\]2.0.CO;2](http://dx.doi.org/10.1130/0016-7606(1972)83[3025:IRGGCN]2.0.CO;2).
- Prantl, H., Brynjolfsson, Skafti, Brynjolfsson, Sveinn, Sailer, R., Stötter, J., 2018. Using Remotely Sensed Data to Identify Perennial Snow Patches in Northern Iceland. Presented at the 5th European Conference on Permafrost, Laboratoire EDYTEM, CNRS, Université Savoie Mont-Blanc, Chamonix, France.
- Prasad, A.M., Iverson, L.R., Liaw, A., 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems* 9, 181–199, <http://dx.doi.org/10.1007/s10021-005-0054-1>.
- Core Team, R., 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Riddle, D.L., Stratford, P.W., 1999. Interpreting validity indexes for diagnostic tests: an illustration using the berg balance test. *Phys. Ther.* 79, 939–948, <http://dx.doi.org/10.1093/ptj/79.10.939>.
- Riley, S.J., DeGloria, S.D., Elliot, R., 1999. A terrain ruggedness that quantifies topographic heterogeneity. *J. Sci.*, 5.
- Rolland, C., 2003. Spatial and seasonal variations of air temperature lapse rates in Alpine regions. *J. Clim.* 16, 1032–1046.
- Rossi, N., Niedrist, G., Asam, S., Tonon, G., Tomelleri, E., Zebisch, M., 2019. A comparison of the signal from diverse optical sensors for monitoring Alpine grassland dynamics. *Remote Sens.* 11, 296, <http://dx.doi.org/10.3390/rs11030296>.
- Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., Harlan, J.C., 1974. Monitoring the Vernal Advancement and Retrogradation (greenwave Effect) of Natural Vegetation (Technical Report No. E75-10354). Texas A&M University; Remote Sensing Center, College Station, Texas.
- Sattler, K., Anderson, B., Mackintosh, A., Norton, K., de Róiste, M., 2016. Estimating permafrost distribution in the Maritime Southern Alps, New Zealand, based on climatic conditions at rock glacier sites. *Front. Earth Sci.* 4, <http://dx.doi.org/10.3389/feart.2016.00004>.
- Schlögel, R., Marchesini, I., Alvioli, M., Reichenbach, P., Rossi, M., Malet, J.-P., 2018. Optimizing landslide susceptibility zonation: effects of DEM spatial resolution and slope unit delineation on logistic regression models. *Geomorphology* 301, 10–20, <http://dx.doi.org/10.1016/j.geomorph.2017.10.018>.
- Schmid, M.-O., Baral, P., Gruber, S., Shahi, S., Shrestha, T., Stumm, D., Wester, P., 2015. Assessment of permafrost distribution maps in the Hindu Kush Himalayan region using rock glaciers mapped in Google Earth. *Cryosphere* 9, 2089–2099, <http://dx.doi.org/10.5194/tc-9-2089-2015>.
- Schratz, P., Muenchow, J., Iturriza, E., Richter, J., Brenning, A., 2019. Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. *Ecol. Modell.* 406, 109–120, <http://dx.doi.org/10.1016/j.ecolmodel.2019.06.002>.
- Scotti, R., Brardinoni, F., Alberti, S., Frattini, P., Crosta, G.B., 2013. A regional inventory of rock glaciers and protalus ramparts in the central Italian Alps. *Geomorphology* 186, 136–149, <http://dx.doi.org/10.1016/j.geomorph.2012.12.028>.
- Seppi, R., Carton, A., Zumiani, M., Dall’Amico, M., Zampedri, G., Rigon, R., 2012. Inventory, distribution and topographic features of rock glaciers in the southern region of the Eastern Italian Alps (Trentino). *Geografia Fisica e Dinamica Quat.* 35, 185–197.
- 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, <http://dx.doi.org/10.1016/j.geomorph.2016.03.015>.
- Steger, S., Kofler, C., 2019. Statistical modeling of landslides: landslide susceptibility and beyond. In: Pourghasemi, H.R., Gokceoglu, C. (Eds.), *Spatial Modeling in GIS and R for Earth and Environmental Science*. Elsevier, Amsterdam, Netherlands, pp. 519–546.
- Strobl, C., Boulesteix, A.-L., Zeileis, A., Hothorn, T., 2007. Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinf.* 21.
- Vapnik, V.N., 2006. Estimation of Dependences Based on Empirical Data, 2nd edn. ed. Information science and statistics. Springer, New York, N.Y., Reprint of 1982 ed. with afterword of 2006. ed.
- Vorpahl, P., Elsenbeer, H., Märker, M., Schröder, B., 2012. How can statistical models help to determine driving factors of landslides? *Ecol. Modell.* 239, 27–39, <http://dx.doi.org/10.1016/j.ecolmodel.2011.12.007>.
- Wahrhaftig, C., Cox, A., 1959. Rock glaciers in the Alaska Range. *Geol. Soc. Am. Bull.* 70, 383, [http://dx.doi.org/10.1130/0016-7606\(1959\)70\[383:RGITAR\]2.0.CO;2](http://dx.doi.org/10.1130/0016-7606(1959)70[383:RGITAR]2.0.CO;2).
- Whalley, W.B., Martin, H.E., 1992. Rock glaciers: II models and mechanisms. *Prog. Phys. Geogr.* 16, 127–186, <http://dx.doi.org/10.1177/030913339201600201>.
- Wood, S., 2019. Package “mgcv”.
Yokoyama, R., Shirasawa, M., Pike, R.J., 2002. Visualizing topography by openness: a new application of image processing to digital elevation models. *Photogramm. Eng. Remote Sens.* 68, 257–266.
- Zischg, A., Mair, V., Lang, K., 2012. PROALP - Kartierung Und Monitoring Von Permafrost in Der Autonomen Provinz Bozen Südtirol, Italien. In: Presented at the INTERPRAEVENT 2012, International Research Society INTERPRAEVEN, Grenoble, France, pp. 421–432.