Geoscience Frontiers 15 (2024) 101822

Contents lists available at ScienceDirect

Geoscience Frontiers

journal homepage: www.elsevier.com/locate/gsf

Research Paper

Adopting the margin of stability for space–time landslide prediction – A data-driven approach for generating spatial dynamic thresholds

Stefan Steger[a,b,*], Mateo Moreno[b,c], Alice Crespi[b], Stefano Luigi Gariano[d], Maria Teresa Brunetti[d], Massimo Melillo[d], Silvia Peruccacci[d], Francesco Marra[e,f], Lotte de Vugt[g], Thomas Zieher[h], Martin Rutzinger[g], Volkmar Mair[i], Massimiliano Pittore[b]

[a] GeoSphere Austria, Vienna, Austria
[b] Center for Climate Change and Transformation, Eurac Research, Bolzano, Italy
[c] Faculty of Geo-information Science and Earth Observation (ITC), University of Twente, Enschede, the Netherlands
[d] Research Institute for Geo-Hydrological Protection, National Research Council (CNR-IRPI), Perugia, Italy
[e] Department of Geosciences, University of Padua, Padua, Italy
[f] Institute of Atmospheric Sciences and Climate, National Research Council (CNR-ISAC), Bologna, Italy
[g] Department of Geography, University of Innsbruck, Innsbruck, Austria
[h] Department of Natural Hazards, Austrian Research Centre for Forests (BFW), Innsbruck, Austria
[i] Office for Geology and Building Materials Testing, Autonomous Province of Bolzano-South Tyrol, Cardano, Italy

ARTICLE INFO

Article history:
Received 27 October 2023
Revised 24 January 2024
Accepted 7 March 2024
Available online 13 March 2024
Handling Editor: Wengang Zhang

Keywords:
Early warning
Space-time model
Rainfall thresholds
Landslide susceptibility, Generalized Additive Mixed Model
Forecasting

ABSTRACT

Shallow landslide initiation typically results from an interplay of dynamic triggering and preparatory conditions along with static predisposition factors. While data-driven methods for assessing landslide susceptibility or for establishing rainfall-triggering thresholds are prevalent, integrating spatio-temporal information for dynamic large-area landslide prediction remains a challenge. The main aim of this research is to generate a dynamic spatial landslide initiation model that operates at a daily scale and explicitly counteracts potential errors in the available landslide data. Unlike previous studies focusing on space–time landslide modelling, it places a strong emphasis on reducing the propagation of landslide data errors into the modelling results, while ensuring interpretable outcomes. It introduces also other noteworthy innovations, such as visualizing the final predictions as dynamic spatial thresholds linked to true positive rates and false alarm rates and by using animations for highlighting its application potential for hindcasting and scenario-building.

The initial step involves the creation of a spatio-temporally representative sample of landslide presence and absence observations for the study area of South Tyrol, Italy (7400 km^2) within well-investigated terrain. Model setup entails integrating landslide controls that operate on various temporal scales through a binomial Generalized Additive Mixed Model. Model relationships are then interpreted based on variable importance and partial effect plots, while predictive performance is evaluated through various cross-validation techniques. Optimal and user-defined probability cutpoints are used to establish quantitative thresholds that reflect both, the true positive rate (correctly predicted landslides) and the false positive rate (precipitation periods misclassified as landslide-inducing conditions). The resulting dynamic maps directly visualize landslide threshold exceedance. The model demonstrates high predictive performance while revealing geomorphologically plausible prediction patterns largely consistent with current process knowledge. Notably, the model also shows that generally drier hillslopes exhibit a greater sensitivity to certain precipitation events than regions adapted to wetter conditions. The practical applicability of the approach is demonstrated in a hindcasting and scenario-building context. In the currently evolving field of space–time landslide modelling, we recommend focusing on data error handling, model interpretability, and geomorphic plausibility, rather than allocating excessive resources to algorithm and case study comparisons.

* Corresponding author at: GeoSphere Austria, Vienna, Austria.
E-mail address: stefan.steger@geosphere.at (S. Steger).

https://doi.org/10.1016/j.gsf.2024.101822
1674-9871/© 2024 China University of Geosciences (Beijing) and Peking University. Published by Elsevier B.V.

1. Introduction

Landslides are geomorphic processes that serve as key sediment sources in mountain areas (Broeckx et al., 2020) while simultaneously posing a risk to infrastructure and human well-being (Glade et al., 2005; Froude and Petley, 2018). Reliable landslide predictions build the foundation for regional early warning and can help to reduce adverse landslide impacts (Piciullo et al., 2018; Guzzetti et al., 2020; Lombardo et al., 2020). The quality of such predictions, however, heavily depends on how well the underlying causes of slope instability can be captured by available data and the employed modelling techniques.

Whether an area is generally prone to shallow slope instability depends on an interplay of predisposing factors, such as topography, subsurface material, vegetation cover, hydro-climatic predisposition and anthropogenic impacts (Crozier, 1989; Glade et al., 2005; Petley, 2010; Schwarz et al., 2012; Bogaard and Greco, 2016; Schmaltz et al., 2017). The timing i.e. temporal occurrence of slope instability is usually determined by highly dynamic weather events, like heavy rainfall or snow melting (Keefer, 2002; Brunetti et al., 2010; Marra et al., 2016; Krøgli et al., 2018; Luna and Korup, 2022). However, a moderate rainfall event may not trigger slope instability in low-susceptibility terrain while creating critical conditions on landslide-prone hillslopes. Thus, to evaluate where and when slope instability may occur requires a joint consideration of various static and dynamic landslide controls (Crozier, 1989).

The Margin of Stability (MoS) represents the conceptual backbone of this research. In contrast to the Factor of Safety (FoS), the MoS relates to the difference between resisting (R) and driving (D) forces acting on a slope (MoS = $R-D$), as opposed to their ratio (FoS = R/D). The MoS is presumed to provide a more straightforward assessment of an area's ability to endure dynamic stressors. For instance, various combinations of R and D (e.g., A: 3 to 2 or B: 6 to 4) can lead to the same FoS value (FoS 1.5), whereas their difference (MoS A: 1; B: 2) may better elucidate the extent of external stress a slope can endure before failing (Crozier, 1989; Glade et al., 2005).

The MoS depends on an interplay of static predisposing factors and the magnitude and intensity of dynamic destabilizing stressors (Fig. 1). Stable terrain is capable to withstand dynamic stressors (i.e., preparatory and triggering factors) to a large degree while marginally stable terrain is more likely to become unstable due to dynamic disturbances. Predisposing landslide controls, such as the topography or geological settings, are assumed to be rather static in time and render a slope more or less prone to landsliding. Dynamic preparatory controls, such as antecedent rainfall or seasonal vegetation effects, modify the MoS in time, but frequently do not initiate slope movement. Instead, triggering factors, such as heavy rainfall, seismic shaking or intense snowmelt, act on comparably short periods and can shift a hillslope from a *"marginally stable"* condition to an *"unstable"* status to initiate a landslide. For instance, the prevailing predisposing factors determine to which degree a slope can withstand dynamic stressors (Fig. 1a vs. 1b). However, comparably *"stable"* terrain can also experience slope instability in response to exceptional dynamic stressors, such as extraordinary preparatory conditions (Fig. 1c) or due to an extreme triggering event (Fig. 1d) (Crozier, 1989; Goudie, 2004; Glade et al., 2005).

Data-driven models are widely used for assessing landslide occurrence over large areas. Landslide susceptibility models put the spotlight on the spatial domain (i.e., predisposing factors in Fig. 1) to map the spatially-varying static propensity of an area to experience slope instabilities (Steger et al., 2016a; Reichenbach et al., 2018; Lima et al., 2022). The resulting maps are often considered a valuable information source for land management and spatial planning (De Graff et al., 2012; Guillard and Zêzere, 2012; Fressard et al., 2014; Petschko et al., 2014). Approaches focusing on the temporal domain (i.e., dynamic factors in Fig. 1) are often concerned with the elaboration of critical triggering conditions. Empirical studies dealing with rainfall-induced landslides primarily focus on the meteorological trigger while usually not explicitly considering confounding effects related to landslide predisposition. For instance, widely adopted empirical rainfall thresholds statistically link landslide occurrence data with associated rainfall to elaborate critical cumulative rainfall amounts and durations (Guzzetti et al., 2007; Brunetti et al., 2010). Comparably few studies explicitly consider the additional effect of preparatory factors related to long-term antecedent precipitation or seasonality (Chleborad, 2000; Guzzetti et al., 2007; Monsieurs et al., 2019b; Luna and Korup, 2022).

Literature shows that the integration of spatial and temporal aspects in large-area landslide prediction is rarely conducted, but selected examples underscore its considerable potential (Lombardo et al., 2020; Collini et al., 2022; Lin et al., 2022; Lombardo and Tanyas, 2022; Ozturk, 2022; Ahmed et al., 2023; Bajni et al., 2023; Moreno et al., 2024). Advanced data-driven models have recently been employed to model landslide occurrence in space and time by explicitly accounting for spatio-temporal stochastic landslide dependencies. However, although these

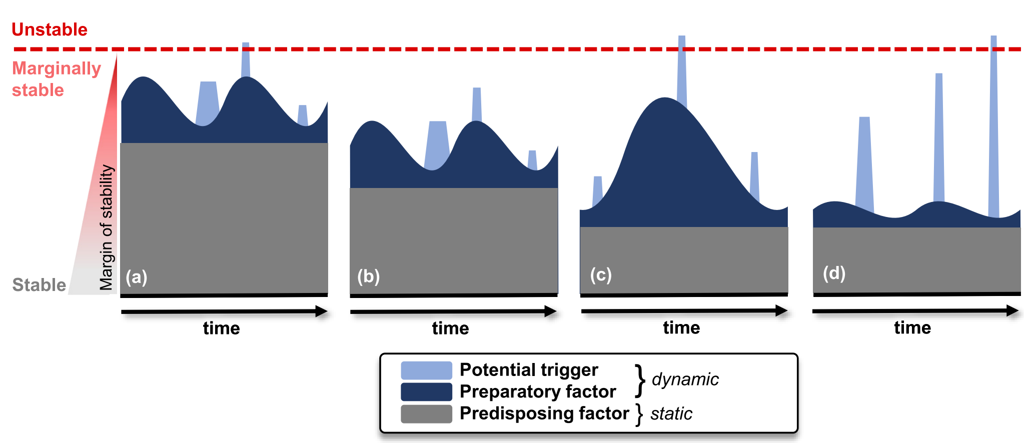

Fig. 1. The Margin of Stability (MoS) depicting factors that control landslide occurrence. Landsliding is initiated ("unstable") due to an interplay of static predisposing factors and dynamic preparatory and triggering factors. The four conceptual cases (a–d) feature different combinations of destabilizing stressors that determine whether a slope will fail (a, c, d) or not (b). Adopted from concepts presented in Crozier (1989) and Zimmermann (1997).

promising studies investigated precipitation-induced landslides, they did not explicitly account for observed precipitation events (Lombardo et al., 2020; Opitz et al., 2022). In the case of rainfall-induced landslides, a joint consideration of landslide predisposition and dynamic landslide controls regularly involves the combination of dynamic components (e.g., meteorological variables) with a pre-existing or separately evaluated landslide susceptibility map through a heuristic combination matrix or a decision-tree approach. However, determining the appropriate weighting of individually evaluated layers during the combination process introduces subjective evaluations (Yang and Adler, 2008; Segoni et al., 2015, 2018; Kirschbaum and Stanley, 2018; Krøgli et al., 2018). Bordoni et al. (2021) used multivariate adaptive regression splines to create a landslide susceptibility model and an additional non-spatial model focusing on hydrometeorological parameters. The final landslide probability index was derived by multiplying the two modelling outputs, which were considered independent. However, assuming independence between temporal and spatial components may impose limitations for landslide prediction (Guzzetti et al., 2006), since landslides and their dynamic controls are inherently spatio-temporal phenomena. A separate consideration might result in biased weights for individual variables and distorted predictions. For instance, a subregion with actual moderate landslide predisposition may nevertheless be characterized by a high number of landslides, because it regularly experiences exceptional precipitation events (cf. Fig. 1d). A conventional susceptibility model does not account for such precipitation dynamics and may, therefore, overestimate the landslide-proneness of this area by assigning too much emphasis (i.e., positive weight) to locally prevailing spatial variables, such as local topographic of soil conditions. In analogy, an empirical study that exclusively considers precipitation variables may overemphasize the significance of relatively low precipitation levels for landslide initiation if a substantial number of landslides were documented in highly landslide-prone terrain, where even moderate precipitation may trigger slope instability (cf. Fig. 1a). Thus, integrating separately evaluated spatial and dynamic components may favour distorted prediction rules.

Recently, data-driven approaches that jointly consider spatial predisposing and measured time-varying dynamic factors have been constructed for two severe landslide events (years 2009 and 2014) that occurred in the Austrian Alpine foreland. The analyses allowed the subsequent assessment of potential changes in landslide occurrence under environmental change (Knevels et al., 2020; Maraun et al., 2022). Inspired by Steger et al. (2023), Ahmed et al. (2023) exploited multi-temporal landslide data related to six rainfall events to create a space–time landslide model for north Vietnam. Nocentini et al. (2023) built a space–time landslide model using a Random Forest classifier and highlighted advantages of high model interpretability. Dynamic landslide prediction models have also been proposed at the global scale (Stanley et al., 2021; Li et al., 2022). The first model created by the National Aeronautics and Space Administration (NASA), which has previously combined separate static and dynamic components, has recently been improved via a fully integrated machine learning approach. This global nowcast model produces a probability-based output that jointly considers spatial and dynamic factors while also allowing the evaluation of trade-offs between wrongly predicted landslides and non-landslides (Stanley et al., 2021).

However, despite these recent advances in dynamic data-driven landslide modelling, comparatively little attention has yet been paid to reduce effects of input data errors on the modelling results. Data quality seems to be less of a concern in data-driven landslide research, even though it is known that a supervised classifier is only as good as its input data, and that available landslide data is usually far from perfect (Guzzetti et al., 2012; Steger et al., 2021). The propagation of data errors into a landslide model should not be ignored, especially since widely-applied quantitative validation procedures have shown limited effectiveness in assessing model quality under data bias conditions (Steger et al., 2016b, 2021). In landslide susceptibility modelling, several researchers took steps to reduce the effects of imperfect input data on model outcomes by mainly modifying the underlying sampling design or by introducing bias-correction procedures during model construction (Steger et al., 2017; Bornaetxea et al., 2018; Jacobs et al., 2020; Lin et al., 2021; Lima et al., 2021). Furthermore, numerous studies focusing on model comparison usually rely on model performance estimates while ignoring aspects related to model interpretability. High model interpretability, however, is necessary for model inference and for conducting plausibility checks (Petschko et al., 2014; Goetz et al., 2015; Steger et al., 2016b, 2021; Lombardo et al., 2020; Collini et al., 2022; Nocentini et al., 2023).

This study delves into space–time landslide modelling by building upon work that derived non-spatial critical precipitation conditions for shallow landslide occurrence (Steger et al., 2023), and by drawing on efforts in susceptibility modelling to address input data errors (Steger et al., 2017, 2021; Bornaetxea et al., 2018). The focus is limited to shallow precipitation-induced landslides, as it should be recognized that other landslide types are controlled by different environmental and triggering conditions. Consequently, modelling other movement-types requires distinctive approaches and considerations (Soeters and van Westen, 1996; Rotigliano et al., 2011; Regmi et al., 2014).

The main aim is to translate the MoS (Fig. 1) into a dynamic spatial landslide initiation model that accounts for errors in the available landslide data while providing interpretable results. Its practical applicability is enhanced by linking the raw predictions to quantitative thresholds, while its potential is demonstrated within hindcasting and scenario construction contexts. The presented data-driven approach stands out as unique by leveraging a 21-year dataset of shallow landslide occurrence in South Tyrol, Italy (7400 km^2), as it:

- puts particular emphasis on model interpretability to provide insights into the interplay of static and dynamic landslide controls.
- accounts for potential spatial and temporal errors in landslide data.
- visualizes the predictions as dynamic spatial thresholds linked to true positive rates and false alarm rates.
- showcases its application potential for hindcasting and forecasting purposes.

2. Study area

South Tyrol is the northernmost province of Italy with an area of 7400 km^2 (Fig. 2). Its general high relief energy is reflected by its altitudinal gradient, which ranges from about 200 m a.s.l. in the southern valley bottoms to about 3900 m a.s.l. in the western lying mountain ranges. The area can be roughly divided into two geological zones separated by the Periadriatic fault. The Southern Limestone Alps with predominant sedimentary rocks in the southeast are opposed by the Central Eastern Alps with predominant metamorphic rocks in the remaining part of the area (Stingl and Mair, 2005). The rural character of South Tyrol is underlined by its land cover: About 45% of the area is covered by forest, often located on the lower lying hillsides, while about 35% is used for agriculture. The climate strongly depends on the altitudinal gradient, with mean annual precipitation ranging from about 600 mm in the lower-lying west to more than 1500 mm in the higher-lying parts in the north and northeast. In general, precipitation amounts are

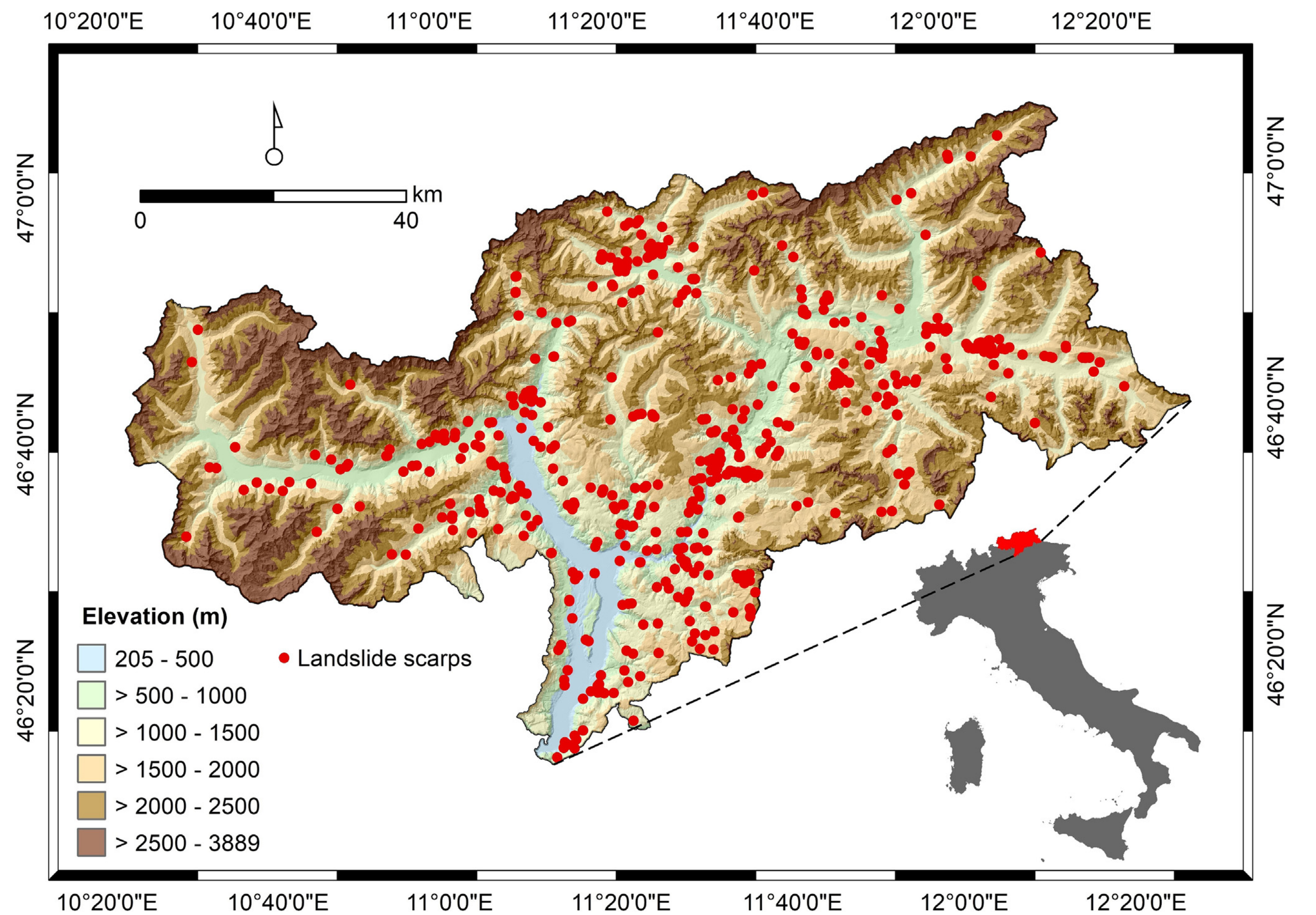

Fig. 2. Overview of the study area. The shown points refer to time-stamped shallow landslide observations that have been registered between 2000 and 2020 in the national inventory called Inventario dei fenomeni franosi in Italia (IFFI). The location of South Tyrol within Italy is shown on the small map.

considerably higher in summer than in winter, when snowfall is common (Crespi et al., 2021).

This research focuses on shallow landslides of the slide-type movement of earth and debris material (Cruden and Varnes, 1996). Available landslide data and associated damage reports indicate that each year a substantial number of events result in damage or affect human activities (Steger et al., 2021). In South Tyrol, most of these landslides are triggered by intense or prolonged precipitation. However, the prevailing topographic conditions, subsurface material, snowmelt, vegetation cover and land use also play a role in determining shallow slope instability (Tasser et al., 2003; Stingl and Mair, 2005; Piacentini et al., 2012; Schlögel et al., 2020; Steger et al., 2021, 2023).

3. Data

3.1. Landslide inventory

This study is based on a 21-year record of precipitation-induced landslides in South Tyrol. The data stems from the South Tyrolean version of the national landslide database IFFI (*Inventario dei fenomeni franosi in Italia*) accessed from the IdroGeo platform (Iadanza et al., 2021). The process to extract day-specific data on precipitation-induced shallow landslides is described in detail in Steger et al. (2023). In summary, the final landslide sample used for modelling consists of 555 precipitation-induced landslide scarp locations with known occurrence dates between 2000 and 2020. Despite being a comprehensive information source for past slope instabilities, this landslide data is known to be biased in space as it systematically reflects damage-causing slope instabilities while underrepresenting landslides in remote and high-altitude locations (Steger et al., 2021).

3.2. Variables representing landslide predisposition

A parsimonious model was ensured by focusing on a limited number of commonly used spatial variables (Reichenbach et al., 2018). Slope morphology variables were derived by the respective SAGA GIS tools (Conrad et al., 2015) from a 30 m Digital Terrain Model (DTM). *Slope angle* and *slope aspect* are amongst the most used variables in landslide susceptibility modelling and were tested for their ability to represent the spatially varying effects of downslope forcing and slope orientation, respectively. The relative topographic position on the hillslope was computed using the SAGA Module "Relative Heights" and its "Normalized Height" output. The resulting variable is henceforth denoted as the *normalized relative height index*. The particular shape/morphometry of local hillslope terrain, which is commonly used as a proxy for surface water and material accumulation, was described via the *convergence index* (van Westen et al., 2008).

Spatial variations of vegetation-related effects are regularly represented by land cover maps (van Westen et al., 2008; Schmaltz et al., 2017). A binary *forest cover* map (forest vs. no-forest) was derived from available land cover information ("Realnutzungskarte Südtirol" version 2015). The general spatial distribution of *lithology* was described via a geological overview map ("Geologische Übersichtskarte Südtirol"). *Mean annual precipitation* from 2000 to 2020 was calculated from daily precipitation grids (Crespi et al., 2021). It should be noted that this climatic variable was not used to describe landslide triggering, but to examine if generally "drier" areas react differently to a certain amount of precipitation compared to areas that frequently experience higher precipitation amounts. In this context, it was also examined whether the results support the postulated landscape equilibrium effect, which suggests that landforms and processes tend to adapt to prevailing climatic conditions (Renwick, 1992).

For modelling, all variables were resampled to the modelling resolution of 30 m $\times$ 30 m to avoid an undesired too detailed description of post-landslide morphology, to reduce noise in the topographic information and to counterbalance positional uncertainties in landslide mapping (Petschko et al., 2014; Steger et al., 2016b). Finally, a categorical variable representing the sampling location (*LOC_ID*) was used as a random effect to account for the spatially nested data structure i.e., the fact that repeated measurements were taken over time at the same location (Zuur et al., 2009).

3.3. Variables representing dynamic landslide controls

Dynamic factors are represented by four types of variables acting at different temporal scales. Gridded daily precipitation data derived from meteorological stations with a cell size of 250 m was used to describe the short-term and medium-term effects of precipitation on landslide occurrence (Crespi et al., 2021). Two types of spatially explicit precipitation variables were tested: triggering precipitation (*T*), representing the cumulative amount of precipitation that fell shortly before and on the observation day and preparatory precipitation (*P*), depicting antecedent cumulative precipitation for a specified time period before *T* (Chleborad et al., 2008). Six candidates representing *T* (day 0 to day 5) and 30 candidates representing *P* (day 1 to day 30) were used to find the best performing *T-P* pair for modelling (cf. Section 4.3). Seasonal effects that cannot be explained by the two precipitation variables were assessed using a day-of-the-year variable (*DOY*). An additional variable (*YEAR*) was used as a random effect to isolate data variability related to a potential bias across years in landslide recording in analogy to Steger et al. (2023).

4. Methods

4.1. Research design

Fig. 3 delineates the methodological framework utilized in this study, which can be categorized into four primary components: (i) Data sampling (green), (ii) Creation of environmental variables (orange), (iii) Variable selection and data-driven modelling (blue), and (iv) Model evaluation and visualization (red). In summary, the first step entailed generating a spatial mask termed the Effectively Surveyed Area (ESA), which restricts data sampling to well-observed and non-trivial terrain, as outlined in Section 4.2. In analogy to the available landslide data, the sampling of landslide absences focused on the 21-year period from 2000 to 2020. The resulting spatio-temporal presence-absence sample was subsequently used to extract associated environmental information, as described in Section 4.3. Modelling built upon a binomial Generalized Additive Mixed Model (GAMM) (Zuur et al., 2009; Pedersen et al., 2019). Predictor variables were selected using a two-step variable selection approach as explained in Section 4.3. The results were assessed in terms of modelled relationships, variable importance, and through multiple cross-validation procedures, as detailed in Sections 4.4 and 4.5. Analysis of the Receiver Operating Characteristic (ROC) curve was conducted to derive three thresholds to create classified maps with categories corresponding to the underlying performance metrics. Dynamic spatial predictions were then generated both in a hindcasting context and for precipitation scenarios (Section 4.5).

4.2. Masking and data sampling

A data-driven spatial landslide model is prone to systematic biases if the underlying study area includes a considerable portion of unvisited terrain, for which landslide absence is typically assumed (Bornaetxea et al., 2018). Subsequent overrepresentation of absence observations within terrain excluded from landslide mappings is likely to result in misleading correlations (Steger et al., 2021). One strategy to address this issue is to narrow the model training to well-observed terrain by limiting the data sampling to the ESA (Bornaetxea et al., 2018). Previous studies defined the ESA by identifying the area visible along a well-investigated path network (Bornaetxea et al., 2018; Knevels et al., 2020). In South Tyrol, the recording of landslides is directly associated with the relevance and proximity to infrastructure while also altitude plays a role. Slope instabilities close to important infrastructure are usually mapped while landslides occurring far from infrastructure or at high altitudes are likely to remain unrecorded. Landslides threatening or damaging crucial infrastructure, such as buildings, highways, or railways, are more likely to be recorded than landslides posing a threat to equally close, but less relevant infrastructure, such as less frequented unpaved roads or trails at high altitudes. Such "data collection effects" were recently integrated into a landslide model for South Tyrol to identify where damage-causing landslides, that are systematically reported, are likely to occur (Steger et al., 2021). In this model, data collection effects were accounted for through various proximity variables, which included distances to key linear infrastructure (roads and railways), and areal infrastructure (buildings) and an interaction term between the distance to less relevant linear infrastructure (pathways) and altitude. Spatially predicting these effects allowed for visualizing the likelihood of landslide recordings, which tended to be highest near critical infrastructure, lower near less significant pathways, especially at higher altitudes and close to zero far from infrastructure. In this study, this raw model output was employed to delineate the ESA (orange area in Fig. 4) by transforming the probability scores into a binary map using the Youden index as explained in Section 4.5. Fig. 4 also visualizes the areas excluded from sampling, namely terrain outside the ESA (blue area) and easy-to-classify non-susceptible terrain, i.e., flat areas below 3.9° for which no-landslide was observed (green area), glacier areas, rocky faces, and water bodies (grey area).

Landslide absence locations were drawn from the sampling area by defining a minimal distance to known landslide locations (red points in Fig. 4) of 150 m. Probability-Proportional-to-Size (PPS) sampling (Singh and Mangat, 1996) was implemented to establish the sampling probability for absence data based on the actual surface area. This approach was used to prevent an underrepresentation of absences in steep terrain, whose actual area is larger than what is represented in a planar perspective (Steger et al., 2021). Between 2000 and 2020, multiple days were randomly selected for each absence location while ensuring a balanced representation across different years and months. Additionally, a minimum temporal gap of 30 days between the observations of the same location was enforced. Finally, all remaining presence and absence observations with precipitation levels of $\leq$ 1.1 mm on the observation day or the day before were excluded from further analysis to focus the model on "wet periods" (Steger et al., 2023).

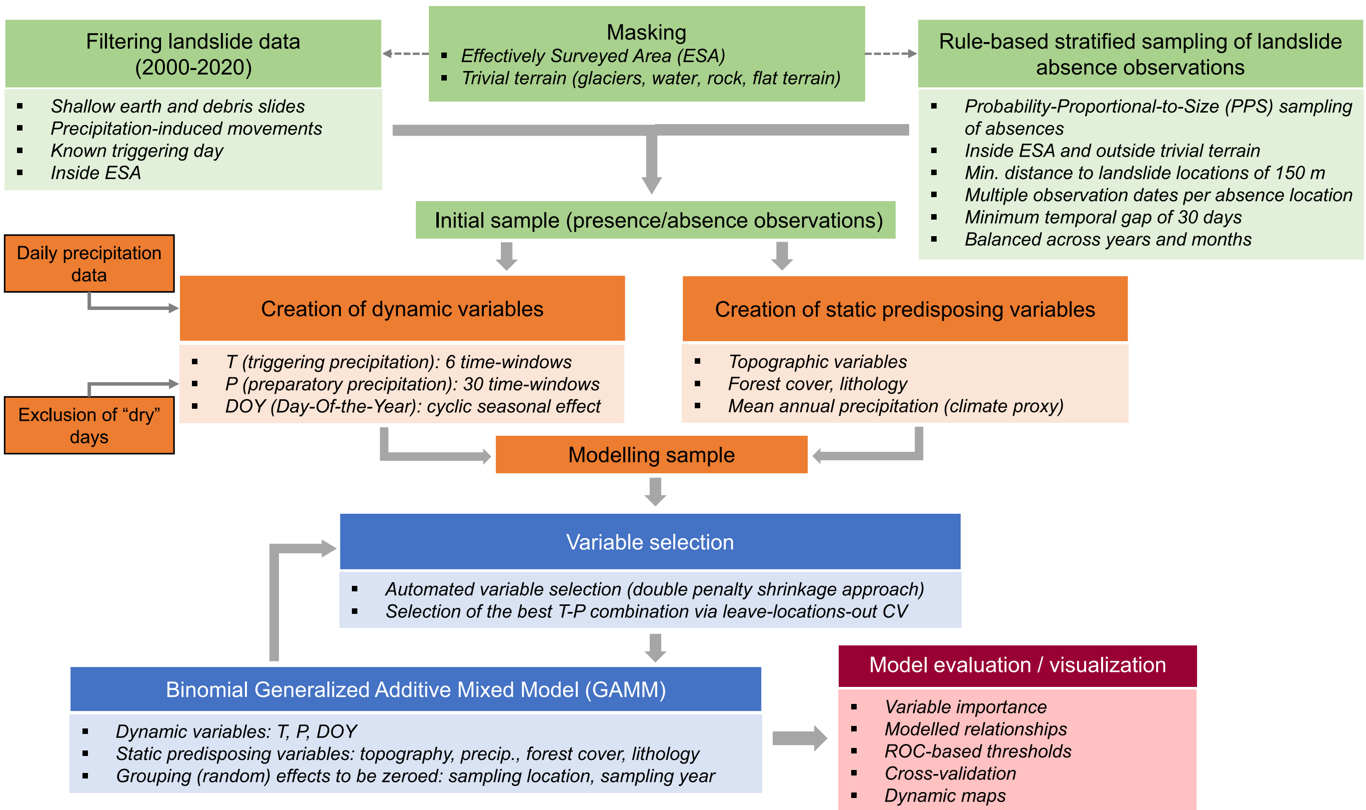

Fig. 3. Overview of the methodical framework of this research.

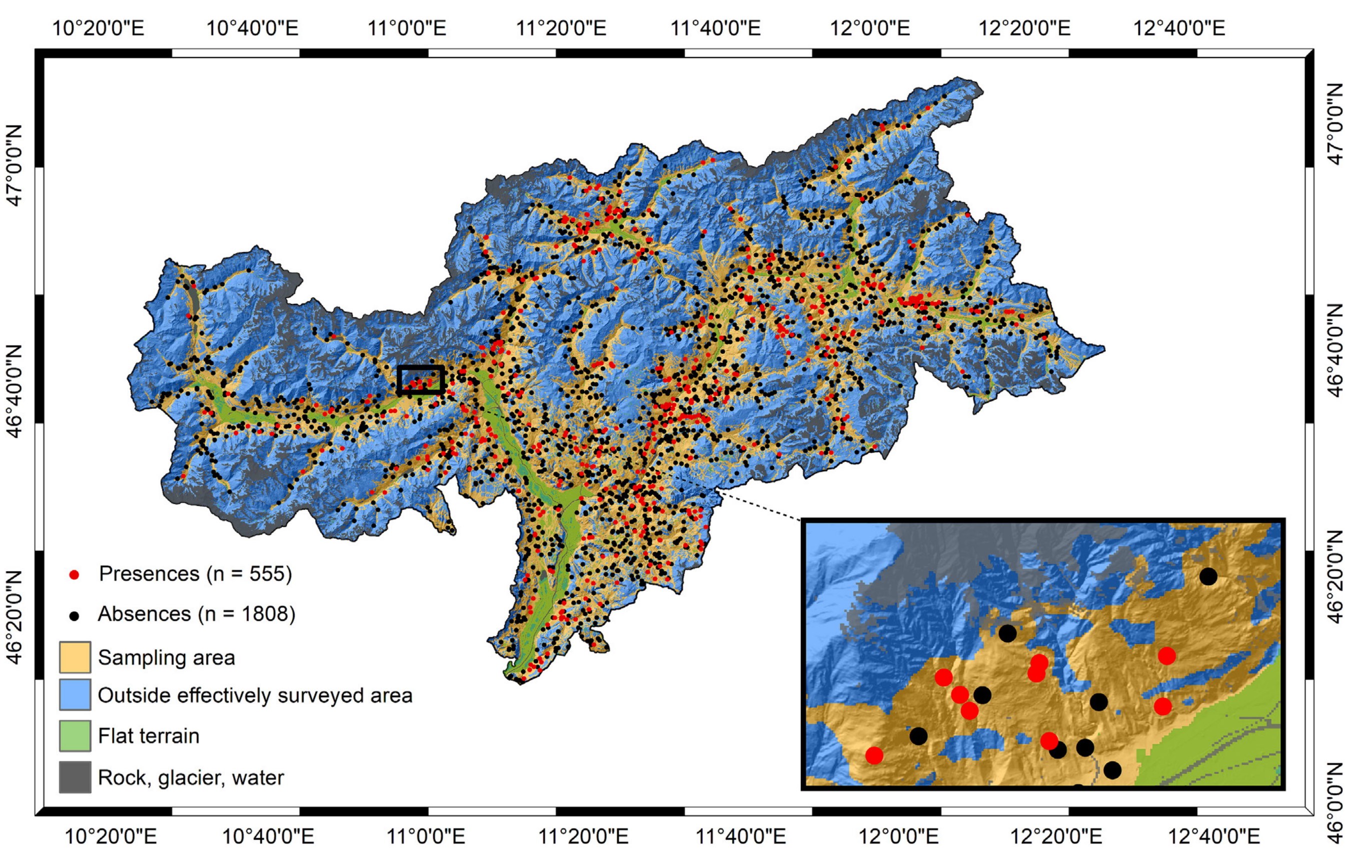

Fig. 4. Visualization of the spatial data sampling design. The area for sampling landslide presence and absence observations is shown in orange along the sampled landslide presence and absence locations. No samples were taken outside the well-observed terrain (blue) or inside trivial and flat terrain (green, grey). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4.3. Generalized Additive Mixed model and variable selection

GAMMs can handle binary response variables, non-linear relationships, interaction effects as well as fixed and random effects, making them well-suited to tackle complex binary classification problems (Zuur et al., 2009; Pedersen et al., 2019). GAMMs have already been used in data-driven landslide modelling to handle spatially or temporally nested data structures and biased landslide data while retaining a high interpretability (Lin et al., 2021; Steger et al., 2023).

The variables included in the GAMM were selected through a two-step iterative procedure. In the first step, all variables described in Section 3.2 and 3.3 were considered when employing a double penalty shrinkage approach, which involves applying a penalty to the range space (i.e., wiggliness penalty restricting the degree of fluctuations) and a second penalty that affects the null space (i.e., shrinkage penalty). This method can be activated by the *select = TRUE* argument in the *mgcv* R package (Marra and Wood, 2011). The initial time-window for representing *T* and *P* precipitation variables was taken from the best performing *T-P* pair of our previous analysis, i.e. 2-day for *T* and 28-day for *P* (Steger et al., 2023).

The second step aimed to assess whether this predefined *T-P* time-window remained the most suitable choice in the current space–time modelling setup. In this context, 25-repeated 5-fold random cross-validation (Section 4.5) combined with a grid-search across 165 *T-P* combinations (180 combination minus the 15 overlapping combinations; cf. grey areas in Fig. 7) was performed. The resultant new best-performing *T-P* pair, which was different from the one identified previously, was then introduced into the first model selection step to determine any potential change in variable selection. This iteration was repeated until a stable variable combination was achieved.

4.4. Evaluation and visualization of modelled relationships

Modelled relationships were evaluated at two levels: at the inter-variable level, through the assessment of variable importance, and at the intra-variable level, by examining partial effects. The variable importance assessment provides insights into the relative contribution of each predictor variable to the response, while the partial effects analysis shows how changing variable values affects modelled landslide occurrence probability.

Variable importance was assessed based on the deviance explained, a well-known measure of the goodness of model fit. In analogy to Goetz et al. (2018), we systematically compared the deviance explained between a full model (all variables) and a series of reduced models, each excluding one predictor or predictor-group. A larger decrease in deviance explained (full model vs. reduced model) indicates a greater relative contribution of the respective variable of interest. Variable importance assessment was conducted for single variables or groups of variables, such as topographic or precipitation variables.

Partial effect plots are graphical representations of the modelled relationship between the response and continuously scaled predictors (Nocentini et al., 2023). Partial effects were visualized as contour plots to visualize the marginal (non-linear) effects of two predictors (i.e., *T* or *P* vs. another predictor) on modelled landslide occurrence in an intuitive way. The relationships shown within these plots depict how the modelled response (i.e., probability of landslide occurrence) changes as the value of the shown predictors varies. Modelled relationships for the categorical variables *forest cover* and *lithology* were interpreted based on odds ratios (ORs). ORs depict the modelled effect size of a certain variable category (e.g., forest) compared to a reference category (e.g., non-forest) to illustrate the relative chance within the respective class to experience landsliding. For instance, an OR of 1.5 would indicate that the estimated likelihood of a landslide occurrence in this particular class (e.g. non-forest) is 1.5 times higher compared to its reference (e.g. forest), which has an OR of 1 (Hosmer and Lemeshow, 2000).

4.5. Calculation of model performance, threshold-building and dynamic maps

The ROC curve was used to classify the prediction maps and to evaluate model performance, also when identifying the best performing *T-P* pair. The ROC is a graphical representation of the performance of a binary classifier as the discrimination threshold is varied. The ROC curve plots, for each possible probability threshold, the associated true positive rate (sensitivity, *y*-axis) against the false alarm rate (1-specificity, *x*-axis). The better a model discriminates between landslide presences and absences, the further its ROC curve moves away from the random classification diagonal towards the top-left corner, increasing the area under the curve. Thus, the Area Under the ROC (AUROC) can be used to summarize the overall model performance of a model. AUROCs usually range from 0.5 (random classification) to 1 (perfect classification) with higher values indicating a better-performing model (Metz, 1978; Hosmer and Lemeshow, 2000).

An AUROC calculated using training data provides insights into the fitting performance of a model. Instead, the predictive performance of a model is usually derived from test data, that was held out during model training. K-fold cross-validation is a standard approach for evaluating the predictive performance of a data-driven model by repeatedly dividing the original data into training data and test data (Schratz et al., 2019). Three cross-validation procedures were implemented based on random, temporal and spatial data partitions.

Random cross-validation utilized 50 random data partitions (i.e., 5 folds, 10 repetitions) to test the predictive performance of the model. In temporal partitioning, data is split based on its temporal component, for example, by using data from several years for training and data from the remaining year for validation. In this study, temporal cross-validation was conducted both, at the month-level and at the year-level by systematically leaving single time units out during model training and testing the model performance for the left-out time periods. Spatial cross validation is based on a spatial partitioning of training and test sets and was conducted on the basis of k-means ($k = 10$) spatial clustering (Brenning, 2012; Schratz et al., 2019) and by dividing the study area into four major geographical zones (i.e. North, West, East, South). Furthermore, final cross-checks with two IFFI-independent landslide inventories were conducted. In this context, we compared the model predictions with landslide information obtained from online newspapers and reports (further called IRPI records; Brunetti et al., 2015; Peruccacci et al., 2023) and from remote-sensing based mappings related to a heavy precipitation event that occurred on the 4th and 5th of August 2016 in the Passeier valley (further called 2016 records; de Vugt et al., 2024).

The ROC was also used to categorize the continuously scaled prediction pattern (i.e., probability scores between 0 and 1) into four classes. Three ROC-derived thresholds, that reflect specific combinations of true positive rate and false alarm rate, were specified for this purpose (Fig. 5). The threshold called *TPR95* relates to a very high true positive rate of 95%. Since almost all (i.e., 95%) observed landslides exceed this threshold, a relatively high false alarm rate can be expected. In contrast, the threshold *TPR25* relates to a low true positive rate of 25% indicating that 75% of all landslides were induced by conditions associated with a lower landslide probability. *TPR25* relates to a comparably high predicted probability score and to a low false alarm rate. The threshold called *OPT* is based on the Youden index and lies between *TPR95* (i.e.,

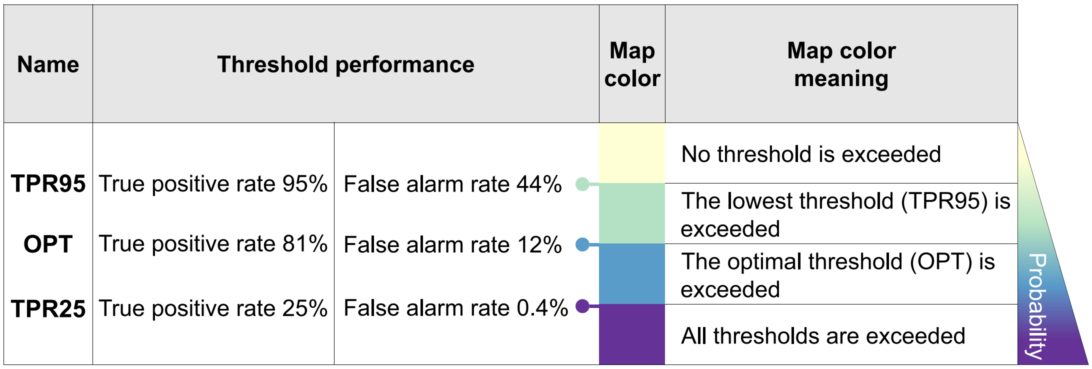

Fig. 5. The thresholds used to visualize the dynamic maps. Utilizing the three thresholds (*TPR95*, *OPT*, *TPR25*) to categorize the probability-based maps results in four distinct classes associated with varying true positive rates (i.e., hit rates) and false alarm rates.

comparably low probabilities) and *TPR25* (i.e., high probabilities). *OPT* balances misclassification rates by maximizing the sum of sensitivity and specificity. It corresponds to the point on the ROC curve with the highest distance from random classification (i.e., the diagonal in the ROC space) and therefore "optimally" separates the two classes (Schisterman et al., 2005). Based on this classification scheme, the final classified map displays four classes, with the optimal threshold (OPT) lying in the middle (Fig. 5).

Finally, the application possibilities of the approach were demonstrated in a hindcasting context and by defining hypothetical precipitation scenarios. Hindcasting aimed to visualize and analyze the spatio-temporal evolution of the landslide thresholds for a severe landslide event that heavily affected the Passeier valley on the 4th and 5th of August 2016 (de Vugt et al., 2024). Such predictions and animations were also created by altering precipitation levels according to spatially uniform hypothetical scenarios. For this purpose, the underlying model predictions were created by a stepwise increase in *T* (5 mm steps) for different fixed antecedent precipitation conditions represented by the *P*-terciles (38 mm: rather dry; 68 mm: average; 104 mm: rather wet).

5. Results

5.1. Sampling results and variable selection

Restricting data sampling to the ESA and non-trivial terrain led to an initial sample consisting of 648 landslide presences (blue bars in Fig. 6a) and 22,693 absences (grey bars in Fig. 6a). This initial sample does not yet consider if the respective observations are related to "*wet periods*" or not. The exclusion of "*dry periods*" resulted in the final modelling sample consisting of 555 landslide presences and 9755 observations not associated with landslide occurrence (Fig. 6b). Thus, around 57% of initially sampled absence observations and 14% of the initially sampled landslide locations were excluded because they were not primarily linked to precipitation. The grey bars in the histogram (Fig. 6b) reflect the general seasonal distribution of precipitation days in South Tyrol, showing that they are more common during the summer. Yet, November has the highest number of landslide observations (blue bars in Fig. 6b).

The two-step variable selection procedure revealed a stable variable combination after two iteration rounds. It led to a variable set consisting of six environmental variables representing landslide predisposition (i.e., *slope angle, normalized relative height, convergence index, forest cover, lithology, mean annual precipitation*), three dynamic variables (i.e., triggering precipitation *T*, preparatory precipitation *P*, seasonal *DOY* effect) and the two grouping effects (i.e., *LOC_ID*, *YEAR*). The best-performing time window for representing *T* and *P* was associated with a mean AUROC of 0.91. This performance was achieved by combining a 2-day cumulative precipitation variable (*Tdays* 1 representing day 0 plus day 1) with an antecedent 26-day time-window (*Pdays* 27) as highlighted in Fig. 7.

5.2. Variable importance and modelled relationships

All selected variables increased the deviance explained by the model (Fig. 8a). The portion of deviance explained was highest for the variable describing short-term precipitation (*T*) followed by *slope angle*, *forest cover* and the variable representing preparatory precipitation (*P*). The remaining variables were associated with a comparatively low relative importance. Since many environmental variables tend to be correlated, a variable importance assessment according to thematic groups was conducted (Fig. 8b). In the current model setup, the dynamic variables were generally assessed to be more important than the static variables (orange bars in Fig. 8b). For instance, leaving out the dynamic precipitation variables *T* and *P* led to a stronger decrease in the deviance explained (0.329) than leaving out all static variables describing landslide predisposition (0.208). In terms of landslide predisposition, topography (0.116) and forest cover (0.108) were estimated to be most influential.

Partial effect plots were used to visualize modelled relationships at the variable level (Fig. 9). Increasing landslide probabilities were generally estimated for increasing amounts of triggering (Fig. 9a–e) and preparatory precipitation (Fig. 9f–j). The trend of this relationship was further modified by the other variables in the model. For example, in case an area experiences 100 mm of *T* or 300 mm of *P*, landslide probabilities were estimated to be highest at ~40° steep terrain (Fig. 9a, f), at concave-shaped landforms (Fig. 9b, g) and on relatively low slope positions (Fig. 9c, h). For a fixed amount of *T* or *P* precipitation, higher landslide probabilities were estimated for generally drier areas compared to their wetter counterparts (Fig. 9d, i). In analogy to the existing non-spatial model (Steger et al., 2023), also the new model setup expressed that more precipitation is required to induce a landslide during summer/autumn, in comparison to winter and spring when vegetation effects and evapotranspiration are less (Fig. 9e, j).

5.3. Model performance

The high fitting performance of the model was confirmed by an AUROC score of 0.91 (Fig. 10). The three selected thresholds were associated with different combinations of true positive rates and false alarm rates. It is shown that the threshold with the highest true positive rate (95.1%; *TPR95*) was associated with a comparably

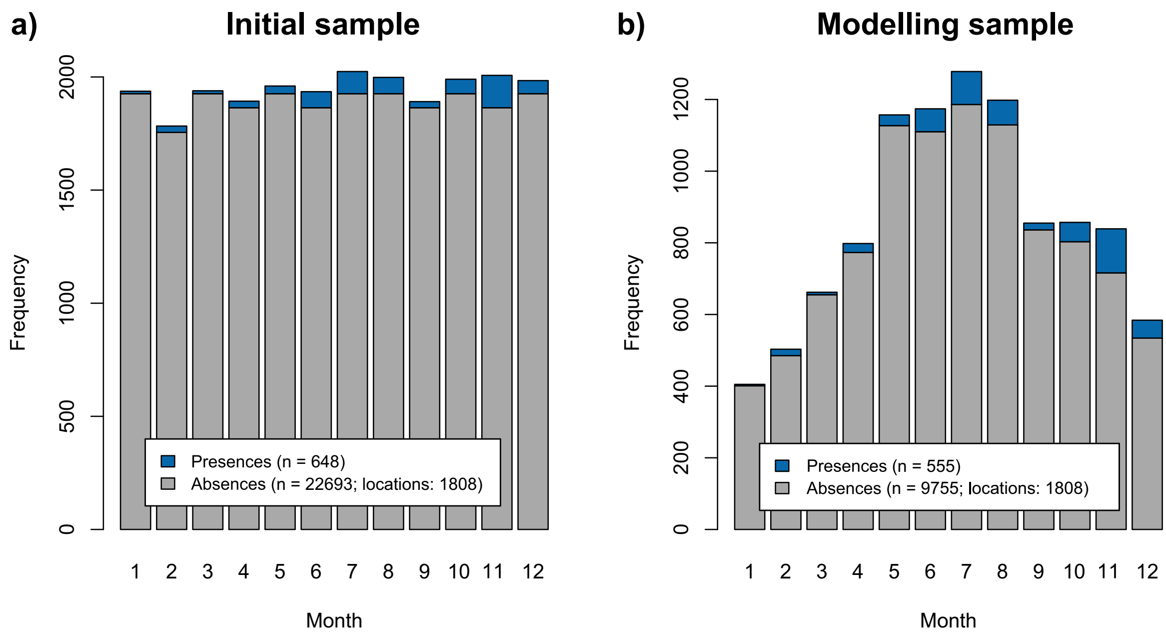

Fig. 6. The data sampling results. The histograms show the frequency of the sampled data before (a) and after (b) applying the precipitation filter. The data used for modelling (b) represents the distribution of "non-dry" landslide presence and absence observations between 2000 and 2020 that are spatially distributed inside the ESA and outside trivial terrain.

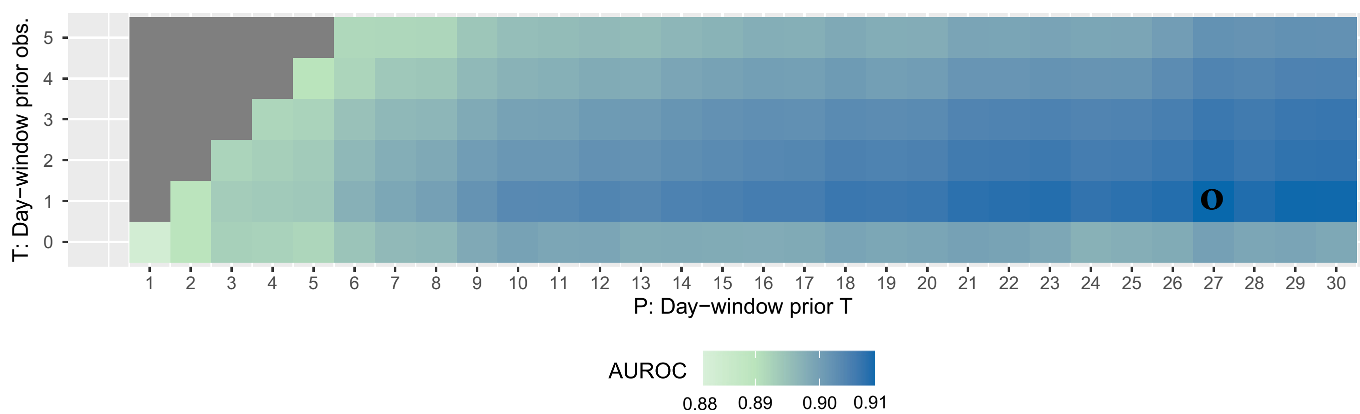

Fig. 7. Results of the grid-search for finding the optimal *T-P* combination. The best-performing time window (black circle) was associated with a mean AUROC of 0.91 and represents the cumulative amount of precipitation for day 0 and day 1 (T_{day} 1) combined with the antecedent 26 days of cumulative precipitation (i.e. P_{day} 27 minus P_{day} 1).

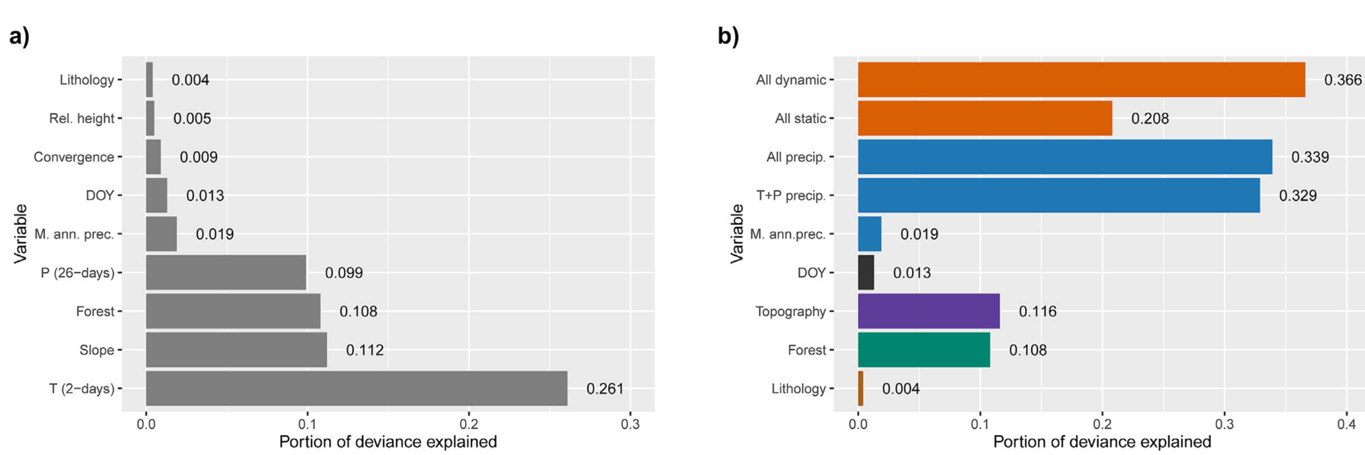

Fig. 8. Variable importance. The relative importance is shown for single variables (a) and variable groups (b) and expressed as the portion of deviance explained within the GAMM.

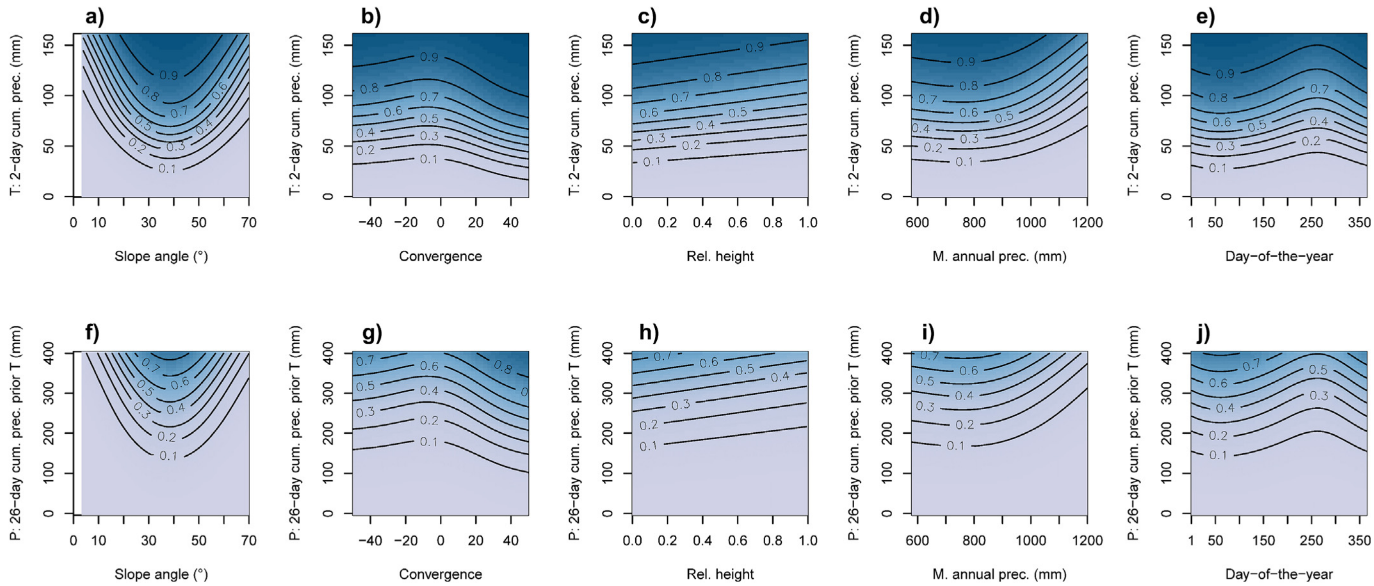

Fig. 9. The modelled relationships. The partial effect plots depict the model predictions by fixing all but the variables in view. The predicted probabilities (contours and color gradient) generally increase with increasing amounts of short-term precipitation (*y*-axis in a–e) and medium-term antecedent precipitation (*y*-axis in f–j). Estimated probabilities also vary across the other continuous variables: slope angle (a, f), convergence (b, g), normalized relative height (c, h), mean annual precipitation (d, i), day of the year (e, j). Odds ratios for the not-shown categorical variables are: Forest cover (Non-Forest 1, Forest 0.13,) and Lithology (Sedimentary 1, Crystalline 0.89, Porphyry 0.88, Plutonite 0.5, Calcschist 0.5).

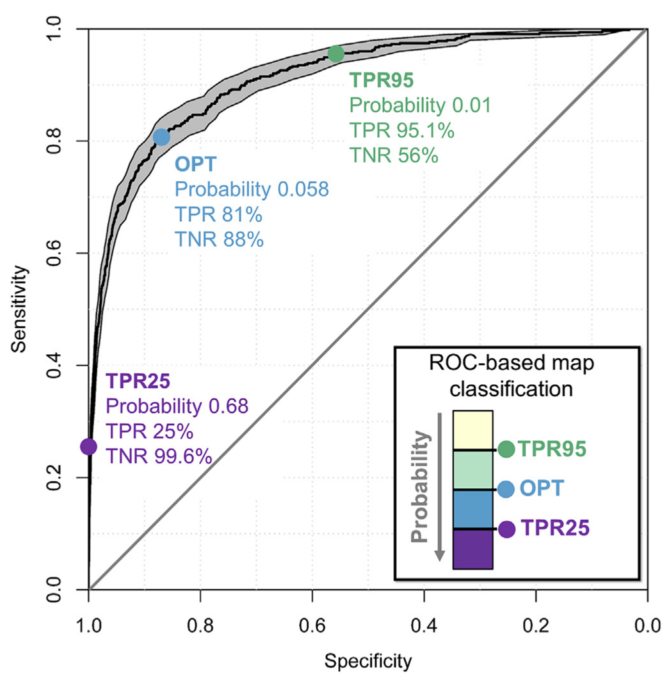

Fig. 10. Fitting performance of the model and metrics related to the three chosen thresholds. The *TPR95* (true positive rate: 95.1%; false alarm rate: 44%) threshold can be linked to the 5% non-exceedance threshold commonly displayed for empirical rainfall thresholds. The *OPT* threshold balances the true positive rate (81%) and the false alarm rate (12%), while *TPR25* (true positive rate: 25%) is only surpassed by a very small number of non-landslide observations (false alarm rate: 0.4%).

high false alarm rate (44%). In contrast, the *TPR25* threshold was associated with a very low false alarm rate (0.4%). The optimal threshold *OPT* was associated with a true positive rate of 81% and a false alarm rate of 12% (Fig. 10).

Random cross-validation revealed a mean AUROC score of 0.91. Temporal cross-validation indicated a relatively high temporal model transferability with a mean AUROC score of 0.89 for the across-year validation and 0.90 for the across-month validation (Fig. 11). While AUROCs were constantly above 0.8 when testing the model against left-out single months (e.g., June as the only month with an AUROC < 0.85), larger deviations were observed for the validation across years. The years 2001 and 2006 stood out with AUROCs below 0.75, while all remaining years were associated with AUROCs above 0.8.

Spatial cross-validation generally confirmed a high spatial transferability of the results with mean AUROCs of 0.9 (Fig. 12). The results indicate that the predictive performance is generally lower in the western region (AUROC of 0.84 in Fig. 12), especially in its far western part (AUROC of 0.77).

The comparison of the spatial predictions with the IFFI-independent landslide datasets showed that all observations surpassed the lowest threshold (*TPR95*). Among the seven IRPI records, five exceeded the optimal threshold. Notably, the 2016 records, which include 55 landslides related to the event in the Passeier valley, exhibited high probability scores. All these observations exceeded the *OPT* threshold, and 87% (n = 48) the highest threshold (*TPR25*), ultimately confirming that this event was considerably more intense than the average landslide-triggering event in South Tyrol.

5.4. Spatial dynamic thresholds

Thresholding of the predictions led to spatially explicit dynamic landslide thresholds (Figs. 13–15). The hindcasting example (Fig. 13) depicts the evolution of the severe precipitation event that triggered more than 50 shallow landslides in the Passeier valley on the 4th and 5th of August 2016 (red circles in Fig. 13a). Associated precipitation maps indicate that this landslide event was induced by exceptionally high and locally restricted short-term precipitation amounts (Fig. 13c) and a moderate quantity of antecedent precipitation (Fig. 13d). The Supplementary Data (Hindcast.gif) visualizes the dynamic evolvement of this event at a daily scale, from 5th of July 2016 to the 15th of August 2016.

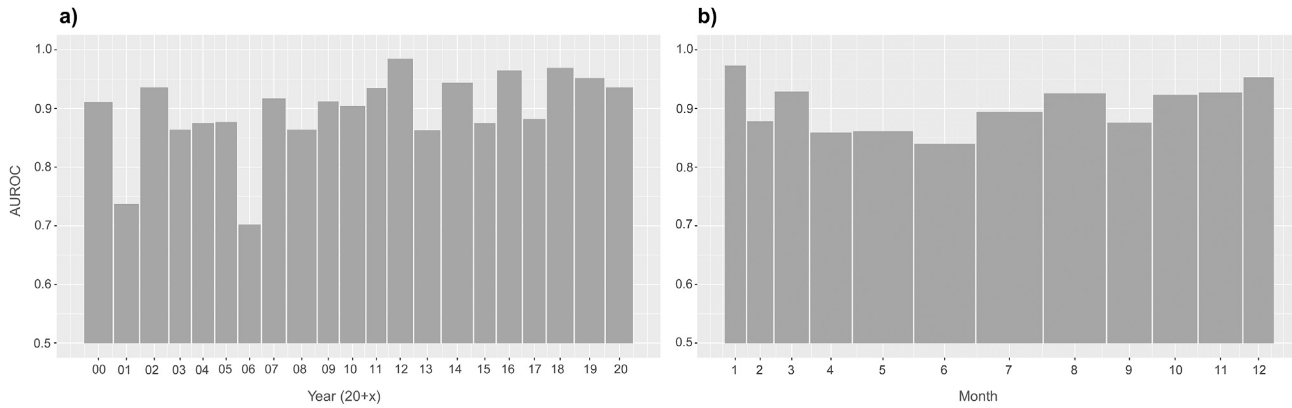

Fig. 11. Temporal cross-validation results. The outcomes are based on leave-one-year-out CV (a) and leave-one-month-out CV (b). Mean AUROC scores were 0.89 (a) and 0.90 (b). Bar widths are proportional to the underlying sample size.

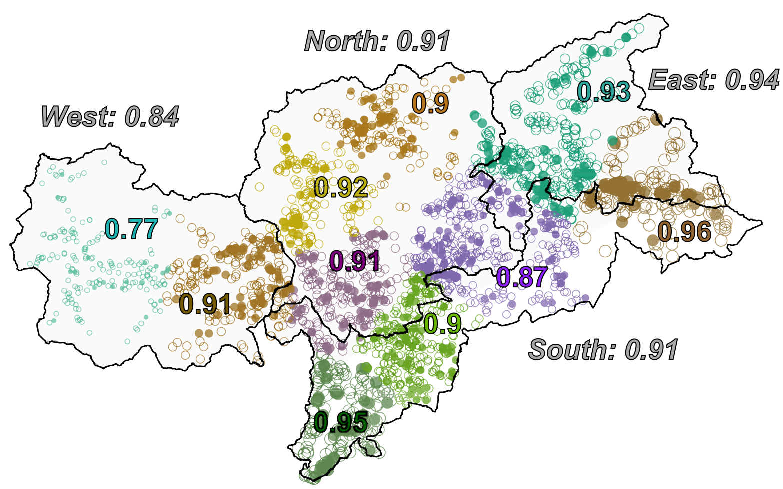

Fig. 12. Cross-validation results based on spatial partitioning of training and test sets. Test set AUROCs of the 10 spatial clusters and associated point locations are shown by different colors (mean AUROC across all clusters: 0.9). Filled circles depict landslide locations, and empty circles depict absence locations, with larger symbols indicating higher AUROCs. The black borders delineate four additional geographical test zones (i.e., North, East, South, West), and associated test set AUROCs (mean AUROC across the four geographical zones: 0.9) are indicated in grey.

Dynamic spatial thresholds were also produced for spatially uniform hypothetical precipitation amounts by systematically increasing short-term precipitation amounts for different fixed antecedent precipitation conditions (Figs. 14 and 15; Scenarios.gif in the Supplementary Data). These maps allowed investigating how much precipitation is required to exceed a certain threshold for a certain area. For instance, in case of higher antecedent precipitation *P* (lower columns in Figs. 14 and 15), less short-term precipitation *T* is required to exceed a specific threshold. The plot depicting the entire region also indicates that the identical amount of short- and medium-term precipitation required to reach a threshold is co-determined by the general dryness/wetness of an area. For example, the typically drier western regions, characterized by lower mean annual precipitation, were observed to surpass a particular threshold earlier in response to a specific precipitation level, in contrast to the generally wetter northern and eastern areas.

A zoom-in to a landslide-prone area in the South (cf. black box on the top-left of Fig. 14) spatially reflects the modelled relationships. It shows the higher dependency of the predictions on changing amounts of short-term triggering precipitation (intra-row comparisons in Fig. 15), compared to varying amounts of preparatory precipitation (intra-column comparisons in Fig. 15). Variation within the same plot depicts the influence of local predisposition on threshold exceedance. All thresholds are exceeded if high precipitation amounts (e.g., T = 100 mm) spatially fall within particularly susceptible terrain.

6. Discussion

This study presents a novel data-driven approach to derive dynamic spatial landslide thresholds by integrating a variety of spatial and dynamic landslide controls. The following discussion embarks on a comprehensive exploration of the key facets of this research. It delves first into measures taken to handle data biases (Section 6.1). Then, important aspects concerning model flexibility, interpretability and visualization are highlighted (Section 6.2). Finally, transferability and application possibilities of the novel approach are discussed (Section 6.3).

6.1. Sampling design and bias handling

Data bias handling was a key consideration in our approach. In the realm of supervised statistical learning, the phrase "*garbage in, garbage out*" highlights the problem that arises when distorted

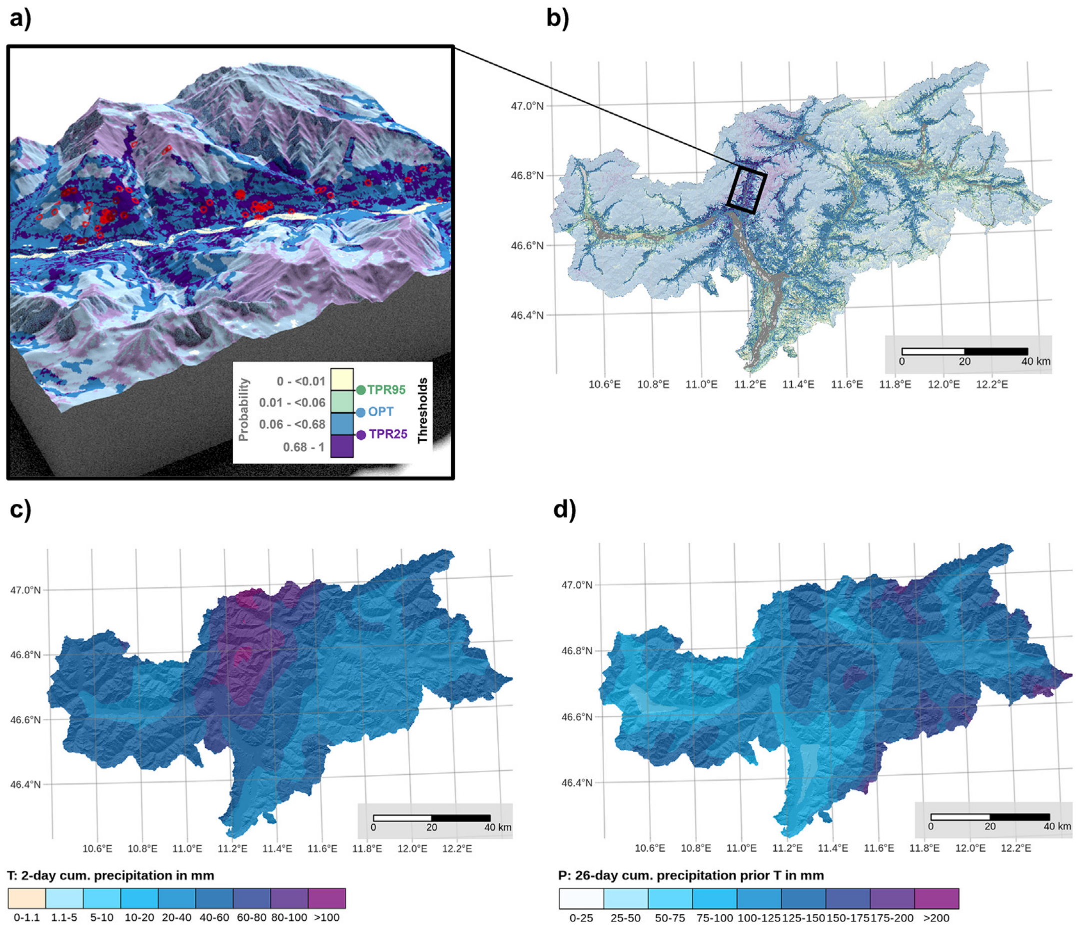

Fig. 13. Hindcasting example for a severe landslide event that occurred on the 4th and 5th of August 2016. The spatial thresholds for the 5th of August are shown for the Passeier valley (a) and for the entire area (b). The 3D map (a) additionally depicts associated landslide observations (red circles) for the section that was most severely hit by the event. The precipitation maps display the associated short-term (c) and antecedent precipitation amounts (d). The animation in the Supplementary Data (Hindcast.gif) visualizes the evolvement of this event from July 5 to August 15. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

modelling outcomes result from flawed or erroneous training data. In machine learning studies, the focus is often on post-data acquisition analysis, and this pattern is also seen in data-driven landslide research, where algorithm comparisons take precedence over evaluating input data quality (Reichenbach et al., 2018).

Landslide data errors that are common for large study sites (Guzzetti et al., 2012) and can compromise the explanatory power of a model, regardless of high model performance scores (Steger et al., 2016b). Therefore, beyond evaluating quantitative performance metrics, it is necessary to obtain insights into potential flaws of the available data and adopt the research design accordingly. For example, to address an underrepresentation of landslide data in specific areas, some studies restrict the data sampling to well-investigated terrain, while others employ mixed-effects modelling to isolate bias-effects (Steger et al., 2017; Bornaetxea et al., 2018; Felsberg et al., 2022). In space–time landslide modelling, systematic data distortions can affect both the spatial and temporal dimension. This study tackled this issue by restricting the spatial data sampling to the ESA while potential across-year inconsistencies in landslide data registration have been isolated using a dedicated year-specific random effect *(YEAR)*. Fostering capabilities of GAMM, an additional spatially explicit random effect (*LOC_ID*) was used to account for temporal data dependencies arising from taking repeated measures over time at the identical location (Zuur et al., 2009; Pedersen et al., 2019).

Another frequently disregarded aspect involves limiting data sampling to areas and time frames that directly relate to the phenomenon of interest. For instance, sampling landslide absences within easily classifiable terrain, such as flat areas, or irrelevant time periods, like dry spells, can lead to a landslide prediction model that primarily distinguishes between floodplains and hillslopes or rainy periods and dry spells, resulting in limited practical usability (Steger and Glade, 2017). This study addressed this issue by exclusively sampling data from non-flat terrain and during non-dry time periods. Also, to effectively capture seasonal effects in landslide occurrence (Luna and Korup, 2022), particular emphasis was placed on creating an absence sample that mirrors the actual distribution of "wet" days across the year, rather than simply sampling an equal number of absences for each month. During the analysis, it became evident that in space–time landslide modelling, rigorous data sampling and bias handling are of great importance.

6.2. Model flexibility, interpretability and visualization

It was aimed to achieve a balance between model flexibility and interpretability. On one hand, the flexibility provided by the GAMM was essential for capturing non-linear relationships (Fig. 9) and for handling hierarchical data structures (Pedersen et al., 2019; Lin et al., 2021). However, it is important to stress that highly flexible classification algorithms tend to sacrifice model

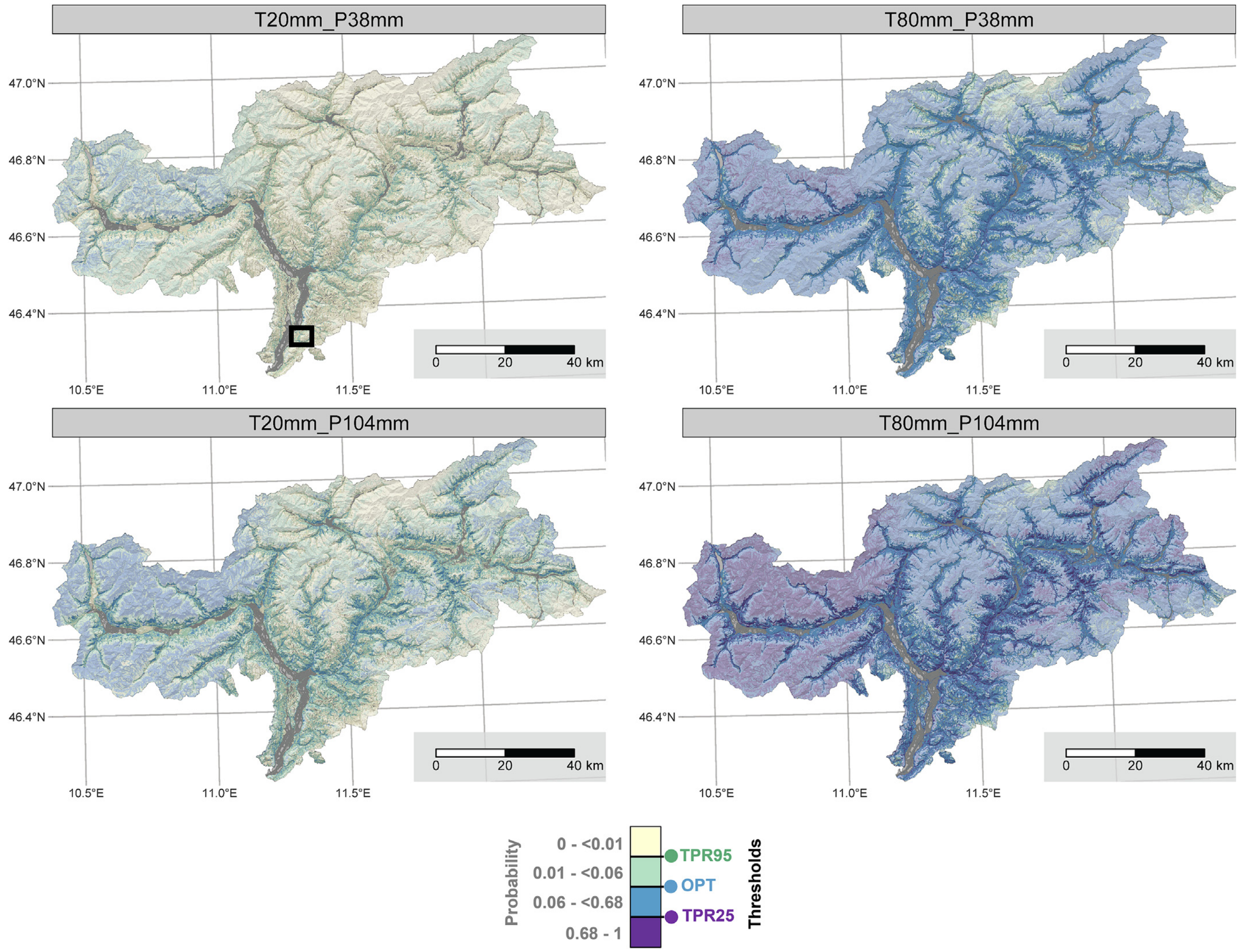

Fig. 14. Example of dynamic spatial thresholds based on spatially uniform hypothetical precipitation for South Tyrol. The predictions are shown for low (T: 20 mm; left column) and high (T: 80 mm, right column) triggering precipitation amounts and for relatively "dry" (25th percentile of P: 38 mm; top row) and "wet" (75th percentile of P: 104 mm; bottom row) antecedent conditions. The black box in a indicates the position of the zoom-in in Fig. 15. The animation in the Supplementary Data (Scenarios.gif) visualizes that in case of higher preparatory precipitation, less short-term precipitation is required to exceed a specific threshold.

interpretability and limit control in shaping the model at a conceptual level, including determining allowed interactions among terms and the complexity of non-linear effects (Goetz et al., 2015). Especially in the presence of common landslide data biases, questioning the reason behind certain modelled relationships is crucial. The importance of having insights into modelled relationships cannot be overstated as it allows exploring whether the modelled relationships plausibly reflect the actual processes under investigation or rather data artifacts (Steger et al., 2016b; Bajni et al., 2023). Once data flaws are addressed to an acceptable level, model interpretability also becomes useful in identifying factors that potentially influence landslide occurrence. This becomes particularly relevant when conveying model outcomes to decision-makers. Explaining what a model does and why it predicts a specific pattern can enhance the acceptance of results (Lombardo et al., 2020; Dahal and Lombardo, 2023).

In the current model setup, dynamic variables emerged as generally more influential than static variables for predicting landslide occurrence in space and time (Fig. 8). This underscored the importance of considering highly dynamic conditions in data-driven landslide prediction whenever possible. The higher relative variable importance of short-term precipitation compared to preparatory precipitation (and mean annual precipitation) is plausible considering the specific landslide type under examination, namely shallow phenomena. However, the highest model performance was obtained by jointly considering static and dynamic variables, confirming that such models can be improved by simultaneously accounting for both, spatially varying landslide predisposition and variables associated to preparatory and triggering factors (Steger et al., 2023).

In analogy to Collini et al. (2022), our model enabled to capture and visualize environmental effects on landslide occurrence operating at various temporal scales. The partial effect plots in Fig. 9 yield valuable insights into the modelled interplay of various environmental landslide controls. For example, an increasing amount of short-term precipitation (T) strongly elevated the probability of landslide initiation. However, in case of a substantial antecedent precipitation amounts (P), the MoS is reduced, and a smaller quantity of short-term precipitation is required to exceed a threshold, as described by Crozier (1989) and Glade et al. (2005) among others. The timing of precipitation throughout the year further influences such effects. In the summer season, characterized by dense vegetation and high evapotranspiration rates, the likelihood of landslides occurring with a particular precipitation level is lower compared to

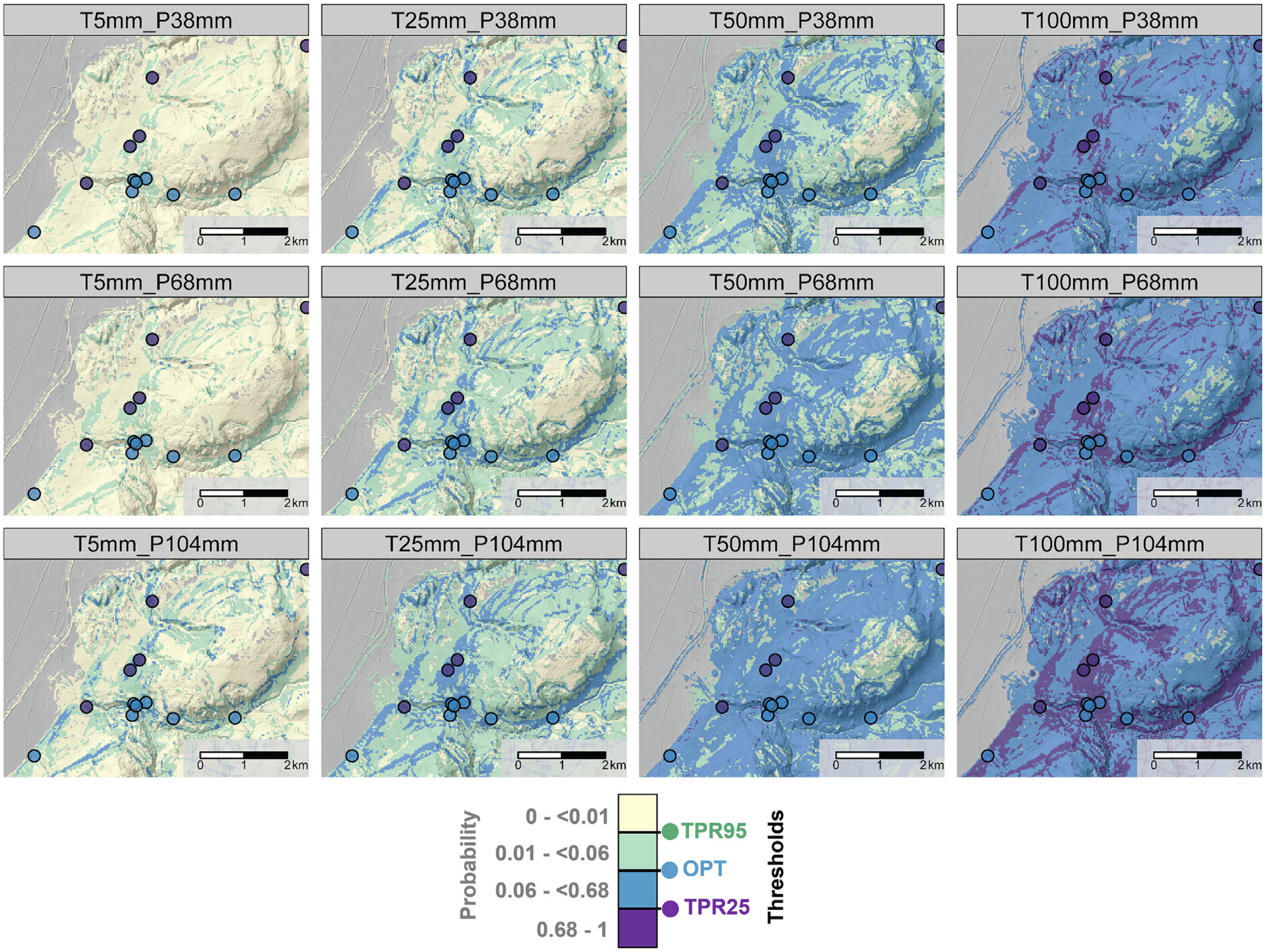

Fig. 15. Scenario example for a landslide-prone subarea of South Tyrol. The shown points depict known landslides while the respective color refers to the respective class-membership at the day of landslide occurrence, i.e., all 12 landslides were triggered above the *OPT* threshold (blue) while 4 of them also exceeded *TPR25* (purple). See also the animation in the Supplementary Data (Scenarios.gif) for further details. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the colder and leafless periods (Steger et al., 2023). Static terrain characteristics also played a role in the model setup. Despite the circular *DOY* variable accounting for seasonal variations, it is anticipated that the ability of the model to capture snow melting is limited. Thus, it is suggested that further model enhancements, particularly in the context of winter and spring predictions, may be achieved through the integration of dynamic snow-related variables (Chleborad et al., 2008; Krøgli et al., 2018).

Steep and unforested terrain, particularly when exhibiting a concave shape that tends to accumulate surface water at lower slope positions, necessitates a comparably low amount of precipitation to reach a specific threshold. This confirmed the added value of considering landslide predisposition in the evaluation of landslide-triggering conditions, especially in areas characterized by varying ground conditions (Peruccacci et al., 2012; Monsieurs et al., 2019a; Nocentini et al., 2023). Interestingly, the incorporation of a spatial proxy accounting for continuous transitions between generally drier and wetter areas (i.e., mean annual precipitation) revealed that, for an identical precipitation event, the model predicted elevated landslide probabilities in regions generally characterized by drier climatic conditions, in contrast to those accustomed to wetter climates. In accordance with the MoS, the model therefore mimics that a particular precipitation event in a dry region can exert an exceptional stress, whereas in a typically wetter environment, the slopes may exhibit higher resilience as they have adjusted to such conditions (Crozier, 1989). Thus, this study presents quantitative evidence supporting the concept of landscape equilibrium, underscoring the tendency of landforms to establish and sustain balance with prevalent climatic conditions (Guidicini and Iwasa, 1977; Renwick, 1992; Aleotti, 2004; Giannecchini, 2006; Giannecchini et al., 2016; Postance et al., 2018).

The preceding section emphasizes that model interpretability plays a pivotal role not only in providing insights into potential factors influencing landslide occurrence, but also for ensuring the plausibility of results. Disregarding the physical meaningfulness of modelled relationships may favor severe errors in subsequent decision-making (Bajni et al., 2023). However, it is also essential to bear in mind that, despite rigorous efforts to minimize errors stemming from input data, a data-driven model does only provide evidence, but cannot prove causation (Steger et al., 2021). Thus, modelled relationships must be interpreted with this caveat in mind.

The visualization of final maps is another important aspect that shapes the usability of the modelling outcomes. Probability-based landslide prediction maps can be visualized in various ways, including displaying raw probability scores or using categorization methods, like equal intervals, natural breaks or quantiles (Hussin et al., 2016; Conoscenti et al., 2016). In decision-making, categorized maps are often preferred because they pretend to clearly differentiate between various "*danger levels*". Nevertheless, assigning practical meaning to a class based on e.g., natural breaks can be challenging. Thus, in the context of spatial planning, classification schemes relying on the portion of observed landslides falling within each class have already demonstrated greater practicality and transparency for end-users (Bell et al., 2013; Petschko et al., 2014). This research showcases the utility of ROC analysis in linking spatial probabilities with true positive rates and false alarm rates, metrics that hold particular relevance in landslide early warning (Leonarduzzi and Molnar, 2020; Stanley et al., 2021). It was shown how the ROC curve can be used to find a quantitative threshold (*OPT*) that balances misclassification rates and optimally separates rather "*untypical*" landslide conditions from rather "*typical*" landslide conditions. For example, if the OPT threshold were to be employed in a binary decision scenario (e.g., issuing a warning or not), it can be stated that 81% of recorded landslides were observed to exceed this cut-point, whereas only 12% of representative wet periods without landslides surpassed it. Given that the predictive performance of the model, with a mean AUROC frequently close to 0.9, closely aligns with its fitting performance (AUROC 0.91), it can reasonably be inferred that these statistics are likely to hold true for future landslide events. The other cut-points, *TPR95* and *TPR25*, illustrated how users can define such thresholds based on their own criteria, such as the portion of landslides exceeding a given threshold. This example also exemplified a well-known limitation when establishing a landslide threshold by ignoring non-landslide information. Namely, it demonstrated that aiming for a very high true positive rate (95%) coincides with a notably high false alarm rate (44%), whereas striving for an exceptionally low false alarm rate (0.4%) is associated with a reduced true positive rate (25%). Consequently, this study advocates for considering non-landslide events when constructing early warning levels (Postance et al., 2018; Leonarduzzi and Molnar, 2020).

6.3. Transferability of the approach and application possibilities

Transferability of research is closely linked to the concepts of generalizability and refers to the degree to which a result can be utilized in different contexts (Coghlan and Brydon-Miller, 2014). In data-driven landslide modelling, random cross-validation is commonly used to test whether a prediction holds true for randomly excluded independent test data. In space–time landslide modelling, dedicated spatial or temporal cross-validation procedures can be considered as means to identify regions or time units in which the model deviates most from the average condition. This in turn can provide further insights into the transferability of the model and systematic model deviations (Steger et al., 2017). Random cross-validation (mean AUROC 0.91) generally confirmed the high capacity of the model to separate landslide conditions from non-landslide conditions at hillslope terrain. Month-related AUROCs were constantly above 0.8, showing a high transferability of the model across seasons. The performance drop observed for the years 2001 and 2006 gave rise to further investigations. A closer look at the associated precipitation amounts showed that those years were associated with a comparably low number of landslide observations (2001: 6; 2006: 14) and precipitation amounts (average *T*: 15 mm for 2001 and 17 mm for 2006; average *P*: 77 mm for 2001 and 44 mm for 2006) similar to those observed for an average non-landslide precipitation event (average *T*: 12 mm; average *P*: 75 mm). Thus, the low AUROCs may reflect the limited effectiveness of precipitation variables to separate presences from absences during those years, leaving the primary discriminatory power to the predisposing conditions.

The methodical approach presented in this research may be exploited within various landslide applications. One prominent example is its utilization in territorial early warning (Piciullo et al., 2018). The versatile approach can also be employed for "*what-if*" scenario exploration, which involves making predictions based on hypothesized precipitation amounts (Fig. 15) or modifications in other variables, such as investigating the effects of deforestation by altering the corresponding forest map. Exploiting meteorological radar estimates may pave the way for landslide nowcasting (Guzzetti et al., 2020) while weather forecasts might be used for landslide forecasting and early warning. In such efforts, however, it has to be taken care that the precipitation data used for prediction does not substantially deviate from the precipitation data used to train the model. Re-calibrating the model or applying bias-correction procedures to align precipitation data (Fustos-Toribio et al., 2022) may be necessary in such cases. Using our approach, the most straightforward way to predict landslide occurrence is to evaluate landslide threshold exceedance using hypothesized spatially uniform precipitation amounts. For example, to determine the exceedance of a landslide threshold two days in advance, the prediction can be made based on the respective *DOY*, observed or estimated cumulative antecedent precipitation amount over the previous 26 days (e.g., *P* 100 mm) along with the precipitation quantity expected for the next two days (e.g., *T* 80 mm), as exemplified within Figs. 14 and 15.

Beyond early warning, the model can also be employed to analyze temporal changes in critical landslide conditions over an extended period. For instance, ongoing research exploits gridded precipitation data (Crespi et al., 2021) to generate more than 14,000 daily prediction maps from 1980 to the present day to provide insights into spatio-temporal trends in critical landslide conditions. By exploring summary statistics for each raster cell, it is also envisaged to create spatially explicit maps that visualize the recurrence interval of raster cells exceeding certain thresholds. In this context, a primary advantage of this approach is that the derived trends are heavily reliant on available historical precipitation data, that might be less influenced by temporal biases compared to long-term records of landslide occurrence. As the model utilizes daily data, another potential application lies in analyzing landslide occurrence under climate change by exploiting emission scenario data (Crozier, 2010). Furthermore, the general methodical framework may be extended to earthquake-induced slope instabilities by replacing the precipitation variables with those describing ground motion patterns. The approach may also be utilized to investigate cascading effects by combining antecedent preparatory events with subsequent landslide triggers. For instance, an option would be to investigate landslide occurrence as a cascading effect of earthquakes followed by rainfall, or vice versa.

While being flexible in terms of the choice of predictor variables, a key requirement for effectively adopting our approach lies in the availability of a sufficiently large dataset that reasonably mimics the real distribution of landsliding over time. To analyze seasonal effects, for instance, a dataset systematically covering the entire year is necessary. Thus, seasonal effects may not be captured adequately when available landslide data pertains only to single or few triggering events. In summary, the minimum sample heavily depends on the complexity of the envisaged model, while different cross-validation procedures (e.g. Steger et al., 2023) may help identify constellations for which the model systematically drops in performance.

7. Conclusion

This research introduces a novel data-driven approach for deriving dynamic, spatially explicit landslide thresholds. It translates the Margin of Stability (Fig. 1) into space–time landslide modelling by integrating landslide controls operating at different temporal scales, including static predisposing factors, seasonal effects, as well as medium-term antecedent and short-term precipitation conditions. In contrast to the limited number of recently published space–time landslide models, the approach not only explores the interplay between static and dynamic landslide controls, but also counteracts the potential propagation of landslide data bias into the results. This was accomplished through a meticulous data sampling and mixed-effects modelling. Emphasis was also placed on generating interpretable outputs through visualizing modelled relationships and by classifying the dynamic spatial predictions according to true positive rates and false alarm rates. Application possibilities are highlighted in a hindcasting context and by generating predictions for hypothetical scenarios involving increased precipitation levels. Outside the field of landslide research, this research may also offer a methodical framework for modelling various types of natural hazards whose occurrence is determined by an interplay of predisposing, preparatory, and triggering factors, such a snow avalanche occurrence or natural wildfire ignition. Within the currently trending domain of data-driven space–time landslide modelling, we promote the prioritization of input data error management, model interpretability, geomorphic plausibility, and result applicability, while avoiding an excessive focus on algorithm comparisons.

CRediT authorship contribution statement

Stefan Steger: Conceptualization, Methodology, Modelling, Validation, Visualization, Writing – original draft. **Mateo Moreno:** Conceptualization, Methodology, Data collection, Visualization, Writing – review & editing. **Alice Crespi:** Conceptualization, Methodology, Data collection, Visualization, Writing – review & editing. **Stefano Luigi Gariano:** Methodology, Writing – review & editing. **Maria Teresa Brunetti:** Methodology, Writing – review & editing. **Massimo Melillo:** Methodology, Writing – review & editing. **Silvia Peruccacci:** Methodology, Writing – review & editing. **Francesco Marra:** Methodology, Writing – review & editing. **Lotte de Vugt:** Conceptualization, Data collection, Writing – review & editing. **Thomas Zieher:** Conceptualization, Data collection, Writing – review & editing. **Martin Rutzinger:** Conceptualization, Data collection, Writing – review & editing. **Volkmar Mair:** Conceptualization, Writing – review & editing. **Massimiliano Pittore:** Conceptualization, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The research leading to these results is related to the PROSLIDE project that received funding from the research program Research Südtirol/Alto Adige 2019 of the Autonomous Province of Bozen/Bolzano – Südtirol/Alto Adige. We also thank the four reviewers for their valuable comments that helped to increase the quality of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gsf.2024.101822.

References

Ahmed, M., Tanyas, H., Huser, R., Dahal, A., Titti, G., Borgatti, L., Francioni, M., Lombardo, L., 2023. Dynamic rainfall-induced landslide susceptibility: a step towards a unified forecasting system. Int. J. Appl. Earth Obs. Geoinf. 125,. https://doi.org/10.1016/j.jag.2023.103593 103593.

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

Bajni, G., Camera, C.A.S., Apuani, T., 2023. A novel dynamic rockfall susceptibility model including precipitation, temperature and snowmelt predictors: a case study in Aosta Valley (northern Italy). Landslides 20, 2131–2154. https://doi.org/10.1007/s10346-023-02091-x.

Bell, Glade, T., Granica, K., Heiss, G., Leopold, P., Petschko, H., Pomaroli, G., Proske, H., Schweigl, J., 2013. Landslide susceptibility maps for spatial planning in Lower Austria. In: Margottini, C., Canuti, P., Sassa, K. (Eds.), Landslide Science and Practice. Springer, Berlin Heidelberg, pp. 467–472. https://doi.org/10.1007/978-3-642-31325-7_60.

Bogaard, T.A., Greco, R., 2016. Landslide hydrology: from hydrology to pore pressure. Wiley Interdiscip. Rev. Water 3, 439–459. https://doi.org/10.1002/wat2.1126.

Bordoni, M., Vivaldi, V., Lucchelli, L., Ciabatta, L., Brocca, L., Galve, J.P., Meisina, C., 2021. Development of a data-driven model for spatial and temporal shallow landslide probability of occurrence at catchment scale. Landslides 18, 1209–1229. https://doi.org/10.1007/s10346-020-01592-3.

Bornaetxea, T., Rossi, M., Marchesini, I., Alvioli, M., 2018. Effective surveyed area and its role in statistical landslide susceptibility assessments. Nat. Hazards Earth Syst. Sci. 18, 2455–2469. https://doi.org/10.5194/nhess-18-2455-2018.

Brenning, A., 2012. Spatial cross-validation and bootstrap for the assessment of prediction rules in remote sensing: the R package sperrorest. In: 2012 IEEE international geoscience and remote sensing symposium, Munich, Germany, pp. 5372–5375.

Broeckx, J., Rossi, M., Lijnen, K., Campforts, B., Poesen, J., Vanmaercke, M., 2020. Landslide mobilization rates: a global analysis and model. Earth-Sci. Rev. 201,. https://doi.org/10.1016/j.earscirev.2019.102972 102972.

Brunetti, M.T., Peruccacci, S., Rossi, M., Luciani, S., Valigi, D., Guzzetti, F., 2010. Rainfall thresholds for the possible occurrence of landslides in Italy. Nat. Hazards Earth Syst. Sci. 10, 447–458.

Brunetti, M.T., Peruccacci, S., Antronico, L., Bartolini, D., Deganutti, A.M., Gariano, S. L., Iovine, G., Luciani, S., Luino, F., Melillo, M., Palladino, M.R., Parise, M., Rossi, M., Turconi, L., Vennari, C., Vessia, G., Viero, A., Guzzetti, F., 2015. Catalogue of rainfall events with shallow landslides and new rainfall thresholds in Italy. In: Lollino, G., Giordan, D., Crosta, G.B., Corominas, J., Azzam, R., Wasowski, J., Sciarra, N. (Eds.), Engineering Geology for Society and Territory –Volume 2: Landslide processes. Springer International Publishing, Cham, pp. 1575–1579. https://doi.org/10.1007/978-3-319-09057-3_280.

Chleborad, A.F., Baum, R.L., Godt, J.W., Powers, P.S., 2008. A prototype system for forecasting landslides in the Seattle, Washington, area. Rev. Eng. Geol. 20, 103–120. https://doi.org/10.1130/2008.4020(06).

Chleborad, A.F., 2000. Preliminary method for anticipating the occurrence of precipitation-induced landslides in Seattle. US Department of the Interior, US Geological Survey, Washington.

Coghlan, D., Brydon-Miller, M., 2014. The SAGE Encyclopedia of action Research. SAGE Publications Ltd. https://doi.org/10.4135/9781446294406.

Collini, E., Palesi, L.A.I., Nesi, P., Pantaleo, G., Nocentini, N., Rosi, A., 2022. Predicting and understanding landslide events with explainable AI. IEEE Access 10, 31175–31189. https://doi.org/10.1109/ACCESS.2022.3158328.

Conoscenti, C., Rotigliano, E., Cama, M., Caraballo-Arias, N.A., Lombardo, L., Agnesi, V., 2016. Exploring the effect of absence selection on landslide susceptibility models: a case study in Sicily, Italy. Geomorphology 261, 222–235. https://doi.org/10.1016/j.geomorph.2016.03.006.

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, 1991–2007. https://doi.org/10.5194/gmd-8-1991-2015.

Crespi, A., Matiu, M., Bertoldi, G., Petitta, M., Zebisch, M., 2021. A high-resolution gridded dataset of daily temperature and precipitation records (1980–2018) for trentino-South Tyrol (north-eastern italian Alps). Earth Syst. Sci. Data 13, 2801–2818. https://doi.org/10.5194/essd-13-2801-2021.

Crozier, M.J., 1989. Landslides: causes, Consequences & Environment. Routledge, New York, London.

Crozier, M.J., 2010. Deciphering the effect of climate change on landslide activity: a review. Geomorphology 124, 260–267. https://doi.org/10.1016/j.geomorph.2010.04.009.

Cruden, D.M., Varnes, D.J., 1996. Landslide types and processes. Transportation Research Board, U.S. National Academy of Sciences, Special Report 247, 36–75.

Dahal, A., Lombardo, L., 2023. Explainable artificial intelligence in geoscience: a glimpse into the future of landslide susceptibility modeling. Comput. Geosci. 176,. https://doi.org/10.1016/j.cageo.2023.105364 105364.

De Graff, J.V., Romesburg, H.C., Ahmad, R., McCalpin, J.P., 2012. Producing landslide-susceptibility maps for regional planning in data-scarce regions. Nat. Hazards 64, 729–749. https://doi.org/10.1007/s11069-012-0267-5.

de Vugt, L., Zieher, T., Schneider-Muntau, B., Moreno, M., Steger, S., Rutzinger, M., 2024. Spatial transferability of the physically-based model TRIGRS using parameter ensembles. Earth Surf. Proc. Land. https://doi.org/10.1002/esp.5770.

Felsberg, A., Poesen, J., Bechtold, M., Vanmaercke, M., De Lannoy, G.J.M., 2022. Estimating global landslide susceptibility and its uncertainty through ensemble modeling. Nat. Hazards Earth Syst. Sci. 22, 3063–3082. https://doi.org/10.5194/nhess-22-3063-2022.

Fressard, M., Thiery, Y., Maquaire, O., 2014. Which data for quantitative landslide susceptibility mapping at operational scale? case study of the pays d'auge plateau hillslopes (Normandy, France). Nat. Hazards Earth Syst. Sci. 14, 569–588. https://doi.org/10.5194/nhess-14-569-2014.

Froude, M.J., Petley, D.N., 2018. Global fatal landslide occurrence from 2004 to 2016. Nat. Hazards Earth Syst. Sci. 18, 2161–2181. https://doi.org/10.5194/nhess-18-2161-2018.

Fustos-Toribio, I., Manque-Roa, N., Vásquez Antipan, D., Hermosilla Sotomayor, M., Letelier Gonzalez, V., 2022. Rainfall-induced landslide early warning system based on corrected mesoscale numerical models: an application for the southern Andes. Nat. Hazards Earth Syst. Sci. 22, 2169–2183. https://doi.org/10.5194/nhess-22-2169-2022.

Giannecchini, R., 2006. Relationship between rainfall and shallow landslides in the southern apuan Alps (Italy). Nat. Hazards Earth Syst. Sci. 6, 357–364. https://doi.org/10.5194/nhess-6-357-2006.

Giannecchini, R., Galanti, Y., D'Amato Avanzi, G., Barsanti, M., 2016. Probabilistic rainfall thresholds for triggering debris flows in a human-modified landscape. Geomorphology 257, 94–107. https://doi.org/10.1016/j.geomorph.2015.12.012.

Glade, T., Anderson, M., Crozier, M.J. (Eds.), 2005. Landslide Hazard and Risk. John Wiley, Chichester, p. 836.

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

Goetz, J.N., Brenning, A., Marcer, M., Bodin, X., 2018. Modeling the precision of structure-from-motion multi-view stereo digital elevation models from repeated close-range aerial surveys. Remote Sens. Environ. 210, 208–216. https://doi.org/10.1016/j.rse.2018.03.013.

Goudie, A. (Ed.), 2004. Encyclopedia of Geomorphology. International Association of Geomorphologists, London, New York, Routledge, p. 2.

Guidicini, G., Iwasa, O.Y., 1977. Tentative correlation between rainfall and landslides in a humid tropical environment. Bull. Int. Assoc. Eng. Geol. 16, 13–20. https://doi.org/10.1007/BF02591434.

Guillard, C., Zêzere, J., 2012. Landslide susceptibility assessment and validation in the framework of municipal planning in Portugal: the case of Loures municipality. Environ. Manage. 50, 721–735. https://doi.org/10.1007/s00267-012-9921-7.

Guzzetti, F., Galli, M., Reichenbach, P., Ardizzone, F., Cardinali, M., 2006. Landslide hazard assessment in the collazzone area, Umbria, Central Italy. Nat. Hazards Earth Syst. Sci. 6, 115–131.

Guzzetti, F., Peruccacci, S., Rossi, M., Stark, C.P., 2007. Rainfall thresholds for the initiation of landslides in central and southern Europe. Meteorol. Atmos. Phys. 98, 239–267. https://doi.org/10.1007/s00703-007-0262-7.

Guzzetti, F., Mondini, A.C., Cardinali, M., Fiorucci, F., Santangelo, M., Chang, K.-T., 2012. Landslide inventory maps: new tools for an old problem: Earth-sci. Rev. 112, 42–66. https://doi.org/10.1016/j.earscirev.2012.02.001.

Guzzetti, F., Gariano, S.L., Peruccacci, S., Brunetti, M.T., Marchesini, I., Rossi, M., Melillo, M., 2020. Geographical landslide early warning systems. Earth-Sci. Rev. 200,. https://doi.org/10.1016/j.earscirev.2019.102973 102973.

Hosmer, D.W., Lemeshow, S., 2000. Applied logistic regression. John Wiley & Sons, New York, p. 397.

Hussin, H.Y., Zumpano, V., Reichenbach, P., Sterlacchini, S., Micu, M., van Westen, C., Bălteanu, D., 2016. Different landslide sampling strategies in a grid-based bi-variate statistical susceptibility model. Geomorphology 253, 508–523. https://doi.org/10.1016/j.geomorph.2015.10.030.

Iadanza, C., Trigila, A., Starace, P., Dragoni, A., Biondo, T., Roccisano, M., 2021. IdroGEO: a collaborative web mapping application based on REST API services and open data on landslides and floods in Italy. ISPRS Int. J. Geo Inf. 10, 89. https://doi.org/10.3390/ijgi10020089.

Jacobs, L., Kervyn, M., Reichenbach, P., Rossi, M., Marchesini, I., Alvioli, M., Dewitte, O., 2020. Regional susceptibility assessments with heterogeneous landslide information: slope unit- vs. pixel-based approach. Geomorphology 356,. https://doi.org/10.1016/j.geomorph.2020.107084 107084.

Keefer, D.K., 2002. Investigating landslides caused by earthquakes – a historical review. Surv. Geophys. 23, 473–510. https://doi.org/10.1023/A:1021274710840.

Kirschbaum, D., Stanley, T., 2018. Satellite-based assessment of rainfall-triggered landslide Hazard for situational Awareness. Earth's Future 6, 505–523. https://doi.org/10.1002/2017EF000715.

Knevels, R., Petschko, H., Proske, H., Leopold, P., Maraun, D., Brenning, A., 2020. Event-based landslide modeling in the Styrian Basin, Austria: accounting for time-varying rainfall and land cover. Geosciences 10, 217. https://doi.org/10.3390/geosciences10060217.

Krøgli, I.K., Devoli, G., Colleuille, H., Boje, S., Sund, M., Engen, I.K., 2018. The norwegian forecasting and warning service for rainfall- and snowmelt-induced landslides. Nat. Hazards Earth Syst. Sci. 18, 1427–1450. https://doi.org/10.5194/nhess-18-1427-2018.

Leonarduzzi, E., Molnar, P., 2020. Deriving rainfall thresholds for landsliding at the regional scale: daily and hourly resolutions, normalisation, and antecedent rainfall. Nat. Hazards Earth Syst. Sci. 20, 2905–2919. https://doi.org/10.5194/nhess-20-2905-2020.

Li, B., Liu, K., Wang, M., He, Q., Jiang, Z., Zhu, W., Qiao, N., 2022. Global dynamic rainfall-induced landslide susceptibility mapping using machine learning. Remote Sens. 14, 5795. https://doi.org/10.3390/rs14225795.

Lima, P., Steger, S., Glade, T., 2021. Counteracting flawed landslide data in statistically based landslide susceptibility modelling for very large areas: a national-scale assessment for Austria. Landslides 18, 3531–3546. https://doi.org/10.1007/s10346-021-01693-7.

Lima, P., Steger, S., Glade, T., Murillo-García, F.G., 2022. Literature review and bibliometric analysis on data-driven assessment of landslide susceptibility. J. Mt. Sci. 19, 1670–1698. https://doi.org/10.1007/s11629-021-7254-9.

Lin, Q., Lima, P., Steger, S., Glade, T., Jiang, T., Zhang, J., Liu, T., Wang, Y., 2021. National-scale data-driven rainfall induced landslide susceptibility mapping for China by accounting for incomplete landslide data. Geosci. Front. 12,. https://doi.org/10.1016/j.gsf.2021.101248 101248.

Lin, Q., Steger, S., Pittore, M., Zhang, J., Wang, L., Jiang, T., Wang, Y., 2022. Evaluation of potential changes in landslide susceptibility and landslide occurrence frequency in China under climate change. Sci. Total Environ. 850,. https://doi.org/10.1016/j.scitotenv.2022.158049 158049.

Lombardo, L., Opitz, T., Ardizzone, F., Guzzetti, F., Huser, R., 2020. Space-time landslide predictive modelling. Earth-Sci. Rev. 209,. https://doi.org/10.1016/j.earscirev.2020.103318 103318.

Lombardo, L., Tanyas, H., 2022. From scenario-based seismic hazard to scenario-based landslide hazard: fast-forwarding to the future via statistical simulations. Stoch. Environ. Res. Risk Assess. 36, 2229–2242.

Luna, L.V., Korup, O., 2022. Seasonal landslide activity lags annual precipitation pattern in the Pacific northwest. Geophys. Res. Lett. 49, e2022. https://doi.org/10.1029/2022GL098506.

Maraun, D., Knevels, R., Mishra, A.N., Truhetz, H., Bevacqua, E., Proske, H., Zappa, G., Brenning, A., Petschko, H., Schaffer, A., Leopold, P., Puxley, B.L., 2022. A severe landslide event in the Alpine foreland under possible future climate and land-use changes. Commun. Earth Environ. 3, 1–11. https://doi.org/10.1038/s43247-022-00408-7.

Marra, F., Nikolopoulos, E.I., Creutin, J.D., Borga, M., 2016. Space–time organization of debris flows-triggering rainfall and its effect on the identification of the rainfall threshold relationship. J. Hydrol. 541, 246–255. https://doi.org/10.1016/j.jhydrol.2015.10.010.

Marra, G., Wood, S.N., 2011. Practical variable selection for generalized additive models. Comput. Stat. Data an. 55, 2372–2387. https://doi.org/10.1016/j.csda.2011.02.004.

Metz, C.E., 1978. Basic principles of ROC analysis. Semin. Nucl. Med. 8, 283–298. https://doi.org/10.1016/S0001-2998(78)80014-2.

Monsieurs, E., Dewitte, O., Demoulin, A., 2019a. A susceptibility-based rainfall threshold approach for landslide occurrence. Nat. Hazards Earth Syst. Sci. 19, 775–789. https://doi.org/10.5194/nhess-19-775-2019.

Monsieurs, E., Dewitte, O., Depicker, A., Demoulin, A., 2019b. Towards a transferable antecedent rainfall—susceptibility threshold approach for landsliding. Water 11, 2202. https://doi.org/10.3390/w11112202.

Moreno, M., Lombardo, L., Crespi, A., Zellner, P.J., Mair, V., Pittore, M., van Westen, C., Steger, S., 2024. Space-time data-driven modeling of precipitation-induced shallow landslides in South Tyrol. Italy. Sci. Total Environ. 912,. https://doi.org/10.1016/j.scitotenv.2023.169166 169166.

Nocentini, N., Rosi, A., Segoni, S., Fanti, R., 2023. Towards landslide space-time forecasting through machine learning: the influence of rainfall parameters and model setting. Front. Earth Sci. 11, 1152130 https://www.frontiersin.org/articles/10.3389/feart.2023.1152130.

Opitz, T., Bakka, H., Huser, R., Lombardo, L., 2022. High-resolution bayesian mapping of landslide hazard with unobserved trigger event. Ann. Appl. Stat. 16, 1653–1675.

Ozturk, U., 2022. Geohazards explained 10: time-dependent landslide susceptibility. Geol. Today 38, 117–120. https://doi.org/10.1111/gto.12391.

Pedersen, E.J., Miller, D.L., Simpson, G.L., Ross, N., 2019. Hierarchical generalized additive models in ecology: an introduction with mgcv. PeerJ 7, e6876.

Peruccacci, S., Brunetti, M.T., Luciani, S., Vennari, C., Guzzetti, F., 2012. Lithological and seasonal control on rainfall thresholds for the possible initiation of landslides in central Italy. Geomorphology 139–140, 79–90. https://doi.org/10.1016/j.geomorph.2011.10.005.

Peruccacci, S., Gariano, S.L., Melillo, M., Solimano, M., Guzzetti, F., Brunetti, M.T., 2023. The ITAlian rainfall-induced LandslIdes CAtalogue, an extensive and accurate spatio-temporal catalogue of rainfall-induced landslides in Italy. Earth Syst. Sci. Data 15, 2863–2877. https://doi.org/10.5194/essd-15-2863-2023.

Petley, D., 2010. Landslide hazards. In: Alcantara, I., Goudie, A. (Eds.), Geomorphological Hazards and Disaster Prevention. Cambridge University Press, pp. 63–74.

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. https://doi.org/10.5194/nhess-14-95-2014.

Piacentini, D., Troiani, F., Soldati, M., Notarnicola, C., Savelli, D., Schneiderbauer, S., Strada, C., 2012. Statistical analysis for assessing shallow-landslide susceptibility in South Tyrol (south-eastern Alps, Italy). Geomorphology 151–152, 196–206. https://doi.org/10.1016/j.geomorph.2012.02.003.

Piciullo, L., Calvello, M., Cepeda, J.M., 2018. Territorial early warning systems for rainfall-induced landslides. Earth-Sci. Rev. 179, 228–247. https://doi.org/10.1016/j.earscirev.2018.02.013.

Postance, B., Hillier, J., Dijkstra, T., Dixon, N., 2018. Comparing threshold definition techniques for rainfall-induced landslides: a national assessment using radar rainfall. Earth Surf. Proc. Land. 43, 553–560. https://doi.org/10.1002/esp.4202.

Regmi, N.R., Giardino, J.R., McDonald, E.V., Vitek, J.D., 2014. A comparison of logistic regression-based models of susceptibility to landslides in western Colorado, USA. Landslides 11, 247–262. https://doi.org/10.1007/s10346-012-0380-2.

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

Renwick, W.H., 1992. Equilibrium, disequilibrium, and nonequilibrium landforms in the landscape. Geomorphology 5, 265–276. https://doi.org/10.1016/0169-555X(92)90008-C.

Rotigliano, E., Agnesi, V., Cappadonia, C., Conoscenti, C., 2011. The role of the diagnostic areas in the assessment of landslide susceptibility models: a test in the sicilian chain. Nat. Hazards 58, 981–999. https://doi.org/10.1007/s11069-010-9708-1.

Schisterman, E.F., Perkins, N.J., Liu, A., Bondell, H., 2005. Optimal cut-point and its corresponding youden index to discriminate individuals using pooled blood samples. Epidemiology 16, 73–81. https://doi.org/10.1097/01.ede.0000147512.81966.ba.

Schlögel, R., Kofler, C., Gariano, S.L., Van Campenhout, J., Plummer, S., 2020. Changes in climate patterns and their association to natural hazard distribution in South Tyrol (Eastern Italian Alps). Sci. Rep. 10, 5022. https://doi.org/10.1038/s41598-020-61615-w.

Schmaltz, E.M., Steger, S., Glade, T., 2017. The influence of forest cover on landslide occurrence explored with spatio-temporal information. Geomorphology 290, 250–264. https://doi.org/10.1016/j.geomorph.2017.04.024.

Schratz, P., Muenchow, J., Iturritxa, E., Richter, J., Brenning, A., 2019. Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. Ecol. Model. 406, 109–120. https://doi.org/10.1016/j.ecolmodel.2019.06.002.

Schwarz, M., Cohen, D., Or, D., 2012. Spatial characterization of root reinforcement at stand scale: theory and case study. Geomorphology 171–172, 190–200. https://doi.org/10.1016/j.geomorph.2012.05.020.

Segoni, S., Lagomarsino, D., Fanti, R., Moretti, S., Casagli, N., 2015. Integration of rainfall thresholds and susceptibility maps in the Emilia Romagna (Italy) regional-scale landslide warning system. Landslides 12, 773–785. https://doi.org/10.1007/s10346-014-0502-0.

Segoni, S., Tofani, V., Rosi, A., Catani, F., Casagli, N., 2018. Combination of rainfall thresholds and susceptibility maps for dynamic landslide hazard assessment at regional scale. Front. Earth Sci. 6, 85. https://doi.org/10.3389/feart.2018.00085.

Singh, R., Mangat, N.S., 1996. Elements of survey sampling. Springer Science & Business Media, p. 412.

Soeters, R., van Westen, C.J., 1996. Slope instability recognition, analysis and zonation. In: Turner, A.K., Schuster, R.L. (Eds.), Landslides: Investigation and Mitigation, Washington, D.C., Transportation Research Board National Research Council, Special Report 247, 129–177.

Stanley, T.A., Kirschbaum, D.B., Benz, G., Emberson, R.A., Amatya, P.M., Medwedeff, W., Clark, M.K., 2021. Data-driven landslide nowcasting at the global scale. Front. Earth Sci. 9, 640043. https://doi.org/10.3389/feart.2021.640043.

Steger, S., Brenning, A., Bell, R., Glade, T., 2016a. The propagation of inventory-based positional errors into statistical landslide susceptibility models. Nat. Hazards Earth Syst. Sci. 16, 2729–2745. https://doi.org/10.5194/nhess-16-2729-2016.

Steger, S., Brenning, A., Bell, R., Petschko, H., Glade, T., 2016b. Exploring discrepancies between quantitative validation results and the geomorphic plausibility of statistical landslide susceptibility maps. Geomorphology 262, 8–23. https://doi.org/10.1016/j.geomorph.2016.03.015.

Steger, S., Brenning, A., Bell, R., Glade, T., 2017. The influence of systematically incomplete shallow landslide inventories on statistical susceptibility models and suggestions for improvements. Landslides 14, 1767–1781. https://doi.org/10.1007/s10346-017-0820-0.

Steger, S., Glade, T., 2017. The challenge of "trivial areas" in statistical landslide susceptibility modelling. In: Mikos, M., Tiwari, B., Yin, Y., Sassa, K. (Eds.), Advancing Culture of Living with Landslides. Springer International Publishing, Cham, pp. 803–808. https://doi.org/10.1007/978-3-319-53498-5_92.

Steger, S., Mair, V., Kofler, C., Pittore, M., Zebisch, M., Schneiderbauer, S., 2021. Correlation does not imply geomorphic causation in data-driven landslide susceptibility modelling – benefits of exploring landslide data collection effects. Sci. Total Environ. 776,. https://doi.org/10.1016/j.scitotenv.2021.145935 145935.

Steger, S., Moreno, M., Crespi, A., Zellner, P.J., Gariano, S.L., Brunetti, M.T., Melillo, M., Peruccacci, S., Marra, F., Kohrs, R., Goetz, J., Mair, V., Pittore, M., 2023. Deciphering seasonal effects of triggering and preparatory precipitation for improved shallow landslide prediction using generalized additive mixed models. Nat. Hazards Earth Syst. Sci. 23, 1483–1506. https://doi.org/10.5194/nhess-23-1483-2023.

Stingl, V., Mair, V., 2005. Einführung in die geologie südtirols: [aus anlass des 32. internationalen geologischen kongresses im Sommer 2004 in Florenz].: autonome provinz bozen-südtirol, amt f. Geologie U. Baustoffprüfung (in German).

Tasser, E., Mader, M., Tappeiner, U., 2003. Effects of land use in alpine grasslands on the probability of landslides. Basic Appl. Ecol. 4, 271–280. https://doi.org/10.1078/1439-1791-00153.

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

Yang, H., Adler, R.F., 2008. Predicting global landslide spatiotemporal distribution: integrating landslide susceptibility zoning techniques and real-time satellite rainfall estimates. Int. J. Sediment Res. 23, 249–257. https://doi.org/10.1016/S1001-6279(08)60022-0.

Zimmermann, M., 1997. Murganggefahr und Klimaänderung - ein GIS-basierter Ansatz. vdf Hochschulverlag AG, 180 p, ISBN: 978-3-7281-2488-3 (in German).

Zuur, A., Ieno, E.N., Walker, N., Saveliev, A.A., Smith, G.M., 2009. Mixed effects models and extensions in ecology with R. Springer, New York, NY, p. 574.