**Pedro Lima · Stefan Steger · Thomas Glade**

# Counteracting flawed landslide data in statistically based landslide susceptibility modelling for very large areas: a national-scale assessment for Austria

**Abstract** The reliability of input data to be used within statistically based landslide susceptibility models usually determines the quality of the resulting maps. For very large territories, landslide susceptibility assessments are commonly built upon spatially incomplete and positionally inaccurate landslide information. The unavailability of flawless input data is contrasted by the need to identify landslide-prone terrain at such spatial scales. Instead of simply ignoring errors in the landslide data, we argue that modellers have to explicitly adopt their modelling design to avoid misleading results. This study examined different modelling strategies to reduce undesirable effects of error-prone landslide inventory data, namely systematic spatial incompleteness and positional inaccuracies. For this purpose, the Austrian territory with its abundant but heterogeneous landslide data was selected as a study site. Conventional modelling practices were compared with alternative modelling designs to elucidate whether an active counterbalancing of flawed landslide information can improve the modelling results. In this context, we compared widely applied logistic regression with an approach that allows minimizing the effects of heterogeneously complete landslide information (i.e. mixed-effects logistic regression). The challenge of positionally inaccurate landslide samples was tackled by elaborating and comparing the models for different terrain representations, namely grid cells, and slope units. The results showed that conventional logistic regression tended to reproduce incompleteness inherent in landslide training data in case the underlying model relied on explanatory variables directly related to the data bias. The adoption of a mixed-effects modelling approach appeared to reduce these undesired effects and led to geomorphologically more coherent spatial predictions. As a consequence of their larger spatial extent, the slope unit–based models were able to better cope with positional inaccuracies of the landslide data compared to their grid-based equals. The presented research demonstrates that in the context of very large area susceptibility modelling (i) ignoring flaws in available landslide data can lead to geomorphically incoherent results despite an apparent high statistical performance and that (ii) landslide data imperfections can actively be diminished by adjusting the research design according to the respective input data imperfections.

**Keywords** Mixed-effects modelling · Slope units · Logistic regression · Landslide inventory · Validation

## Introduction

In the last decades, there has been an increase in the reporting of landslide phenomena that caused damage or threatened society (Petley, 2012). Recent reviews highlight a growing number of publications dealing with the quantitative spatial prediction of landslides (Reichenbach et al., 2018; Steger & Kofler, 2019; Wu et al., 2015). For very large areas, such assessments are usually based on statistical methods and include national-scale analyses (e.g. (Dikau & Glade, 2003; Domínguez-Cuesta & Bobrowsky, 2017; Ferentinou & Chalkias, 2013; Gaprindashvili & Westen, 2015; Graff et al., 2012; Komac, 2006; Komac & Ribicic, 2006; Liu et al., 2013; Malet et al., 2008; Sabatakakis et al., 2013; Trigila et al., 2013)), continental-scale assessments (e.g. (van Den Eeckhaut et al., 2012; Günther et al., 2013; Günther et al., 2014; Wilde et al., 2018)) and global models (Hong et al., 2007; Lin et al., 2017; Nadim et al., 2006). Lack of reliable spatial geotechnical information and restrictions related to computational resources are well known to hamper the application of physically based spatial landslide models for such large territories (Corominas et al., 2014).

The primary purpose of a landslide susceptibility map is to spatially depict the relative likelihood of landslides affecting an area under a given set of geo-environmental conditions (Brabb, 1984; Fell et al., 2008; Glade & Crozier, 2005). The selection of an appropriate modelling approach should consider the size and characteristics of the target area and the availability and quality of input data (Corominas et al., 2014; Guzzetti, 2005). Under real-world data conditions, the size of a study area is likely to influence the consistency and quality of available geo-environmental and landslide information. Compared to physically based slope stability models, statistical landslide susceptibility analyses are more flexible in terms of input data and thus often applied for the assessment of large areas (Cascini, 2008; Corominas et al., 2014; Sabatakakis et al., 2013; van Westen et al., 2008). Statistically based landslide susceptibility assessments are based on the assumption that the potential location of an upcoming slope failure can be estimated by analysing past landslides and their relation to spatial geo-environmental variables. In most cases, the underlying classification algorithms build a statistical relation between past landslide presence/absence information and a set of explanatory variables that describe the underlying terrain conditions, e.g. the topography, lithology and land cover. The ensuing prediction rule often expressed as a susceptibility score between zero and one is then applied to each terrain unit of a study area (e.g. raster cell, slope unit) (Jacobs et al., 2020). For validation, the predicted susceptibility score is confronted with model-independent landslide observations in order to get insights into the prediction capability of the model (Chung & Fabbri, 2003; Frattini et al., 2010; Reichenbach et al., 2018). Several studies have shown that the explanatory power of statistically based spatial landslide predictions is particularly dependent on the reliability of landslide training data (Ardizzone et al., 2002; Harp et al., 2011; Steger et al., 2016a; Zêzere et al., 2017).

In its basic form, a landslide inventory represents a collection of past landslide locations. For most scientific purposes, a further distinction between landslide types is required (Corominas et al., 2014; Guzzetti et al., 2012). Additional information on the spatial accuracy of events, their temporal occurrence and mapping

uncertainties can further enhance the usability of inventory data (Guzzetti et al., 2012). However, for very large areas, consistent and accurate landslide information is rarely available and the creation of an assembled landslide inventory based on a variety of sources is common practice (Malamud et al., 2004). The resultant spatially heterogeneous data quality poses a challenge for the spatial prediction of landslide phenomena using statistically based approaches (Ardizzone et al., 2002; Fressard et al., 2014; Harp et al., 2011; Hussin et al., 2016; Steger et al., 2017).

The positional accuracy of a landslide inventory is known to be associated with the underlying data source and mapping technique applied (Guzzetti et al., 2012). To what extent positional inaccuracies are propagated into the final modelling and validation results also depends on the chosen raster resolution and the type of terrain representation (Jacobs et al., 2020; Steger et al., 2016b; Steger et al., 2020). Alternatives to a raster-based representation of the landscape, like slope units (Alvioli et al., 2016), are gaining increasing popularity in the context of statistical landslide susceptibility modelling, also because they are considered to be geomorphically more meaningful and potentially less sensitive to landslide positional errors and noise in geo-environmental input data (Alvioli et al., 2016; Camilo et al., 2017; van Den Eeckhaut et al., 2009; Guzzetti et al., 2006; Jacobs et al., 2020; Schlögel et al., 2018).

Another concern in statistical landslide susceptibility modelling is the spatial representativeness (i.e. completeness) of available landslide information (Guzzetti et al., 2012; Steger et al., 2017). A common source of such spatial bias relates to the underreporting of landslides that did not cause social, infrastructural or monetary damage leading to an underrepresentation of landslides far from settlements and infrastructure (Bell et al., 2012; Brardinoni et al., 2003; Petschko et al., 2014; Sabatakakis et al., 2013). Human activity (e.g. earthmoving) and natural processes (e.g. erosion or ecological succession) may as well induce a spatially changing completeness of landslide information among different land cover types (Bell et al., 2012; Petschko et al., 2016). Varying completeness of landslide data can also be related to the boundaries of administrative units. For example, the resources allocated for landslide mapping may vary among administrative units, such as political counties, provinces or municipalities, which can introduce another type of spatial bias (Steger et al., 2017; Trigila et al., 2013). Merging existing landslide inventories from different study sites may also introduce spatial inconsistencies because areas, where less detailed mapping campaigns were performed, are likely to under represent the portion of past landslide occurrences.

Literature indicates that the positional accuracy and spatial consistency (i.e. representativeness) of mapped events are frequently ignored in statistical susceptibility modelling. Especially for large study sites, where data inconsistencies represent the rule rather than an exception, flawed landslide information may lead to erroneous landslide susceptibility models, which in turn may negatively influence the explanatory power of each subsequent decision or analysis. Previous publications highlighted some strategies that may counteract associated error propagations by focusing on the sampling of non-landslide locations (van Den Eeckhaut et al., 2012), positional inaccuracies of landslide information (Jacobs et al., 2020; Steger et al., 2016b) and systematic inventory-based incompleteness (Steger et al., 2017).

This research aims (i) to assess landslide susceptibility for the entire Austrian territory by (ii) counterbalancing landslide inventory-based incompleteness and by (iii) minimizing the effect of positional inaccuracies through a slope unit terrain representation. The application of different classification algorithms (i.e. logistic regression vs. mixed-effects logistic regression) and different terrain representations (i.e. grid-based vs. slope units) allowed the comparison and evaluation of four models in terms of their ability to counterbalance flawed landslide information. Only shallow translational earth and debris slides (further termed landslides) were considered within this analysis (Cruden & Varnes, 1996; Dikau et al., 1996).

## Study area

Containing approximately 84,000 km², the territory of Austria (Fig. 1) is located centrally on the European continent and its approx. 8.8 million inhabitants are distributed across nine federal provinces (Statistik Austria, 2017). The topography of Austria can be described as undulating to flat in the East and predominantly mountainous in the West. The Danube river basin in the East crosses the political capital Vienna in the very East. In contrast, the alpine landscape in the central and western parts exhibits elevations of up to 3798 m asl. at the Großglockner peak. From the north-eastern lower and gentle terrain to the high alpine landscape in the central and western parts, the Austrian territory is characterized by a considerable morphological variety. Landslides represent a substantial threat to private and public properties, critical infrastructure and the population where the landslide favouring geomorphological conditions coincides with specific socio-economical and demographical settings (Petschko et al., 2013; Schweigl & Hervás, 2009).

The geological setting of the Austrian territory is defined by a Northern part mainly dominated by the presence of the Bohemian Massif and Molasse zones. These partly deforested flat to hilly landscapes are located within the northern part of Austria (Krenmayr et al., 2000). The Helvetic and Flysch zones represent relatively narrow lithological units located south of the Molasse zone and correspond to sedimentary deposits. The Flysch zone, an elongated sandstone-rich unit, represents a gentle to hilly landscape that originated from the depositional processes dating from Upper Cretaceous and Neogene. The Flysch zone is well known to be particularly prone to landslides of slide-type movement (Petschko et al., 2014; Schwenk, 1992; Terhorst & Damm, 2009). Various lithologies like the Greywacke zone, the Calcareous Alps, the Penninic unit, crystalline rocks and the Periadriatic compose the high elevation alpine terrain, where landsliding is common, but often underestimated (Krenmayr et al., 2000).

The climate is influenced by the Alpine territory, which represents 62.8% of the total area of Austria (BMLFUW, 2007), with strong influences from the Atlantic continental, sub-Mediterranean and polar or subpolar air masses. Temperature regimes largely depend on elevation ranges and seasonal trends (Auer et al., 2007; Hiebl & Frei, 2016). The mean annual rainfall is also greatly variable over the territory. The humid season also corresponds with the warmer period. Precipitation average annual values range from a few hundred millimetres per year (400 to 600 mm/year in the Eastern part) to more than 2000 mm/year in some of the Alpine areas (BMLFUW, 2007). Localized, highly intense rainfall events are mainly concentrated during the
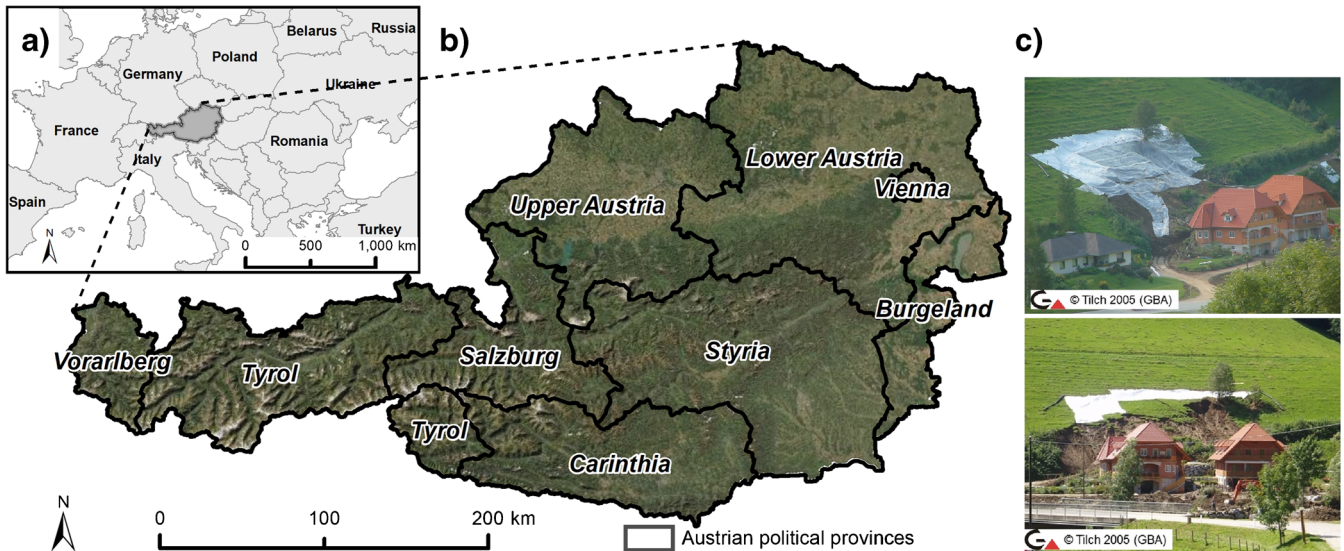
**Fig. 1** (A) The location of Austria at the central European territory. (B) Overview of the Austrian morphological and political territory. (C) Shallow landslide events that occurred within the Austrian territory. Photographs were provided by the Geological Survey of Austria

summer-autumn period and can be manifested as heavy thunderstorms, bringing heavy hail, in summer, occurring throughout the country. However, comprehensive historical data analysis has shown significant regional and different seasonal tendencies (Auer et al., 2007).

The predisposition for landsliding in Austria depends on an interplay of different environmental factors, such as lithology and its weathering products, topography, land cover and human impact (Bell et al., 2012; Schweigl & Hervás, 2009). Extreme precipitation events, as well as rapid snow-melting, are known to represent typical landslide triggering factors (Glade et al., 2020). Severe landslide events caused major consequences, for instance in August 2005 in Gasen und Haslau, where hundreds of landslides were reported after a heavy rainfall of 200 mm in 48 h (Tilch, 2009). In June 2009, thousands of landslides were triggered by heavy rainfalls in Feldbach (Hornich & Adelwöhrer, 2010).

### Data

#### Landslide inventory

The landslide information used for this study consists of 23,891 shallow translational landslides (Cruden & Varnes, 1996; Dikau et al., 1996; Hungr et al., 2014). The final landslide inventory was compiled from nine sub-inventories, which differ in their spatial reference area (i.e. mapping domain) and applied mapping technique. From now on, the term "mapping domain" is used to describe the spatial extension (i.e. a polygon) delimiting the area of each sub-inventory. Complementary inventory characteristics, such as the coverage area (km²), density (slides/km²), the mapping technique used, location description and also a quality indicator about the positional accuracy of each inventory, are described within Table 1 and represented by Fig. 2(D). Although information on the trigger mechanism was not available for all the sub-inventories, it is known that commonly the main triggers of shallow earth and debris slides in

Austria are associated with hydro-meteorological events and not of tectonic origin (Schweigl & Hervás, 2009). Consequently, the present study refers to hydro-meteorological landslide trigger only.

Available landslide inventories are a more or less accurate representation of the past landslide locations (Guzzetti et al., 1999; Guzzetti et al., 2012). Detailed new mapping campaigns that cover very large territories (e.g. entire nations) are costly and time-consuming and do not necessarily lead to a representative landslide data set since the footprint of old landslides might be absent within specific areas. In practice, it usually remains to the researcher to work with the best available data by keeping in mind its limitations. For the present study, the spatial bias in the available landslide information plays a significant role. Literature suggests that in Austria, specific land cover features may over- or underrepresent the true landslide occurrence (Bell et al., 2012; Petschko et al., 2014). Although the compiled inventory contains a considerable number of landslides, the first inspection of available meta-data and the available spatial information indicated heterogeneous completeness among the different mapping domains (Fig. 2(D)). For instance, the very low landslide density in the mountainous region of Tyrol is likely associated with the absence or unavailability of a detailed and province-specific landslide inventory (i.e. only Mp1 was available). In contrast, the comparably high landslide density observed for the province of Lower Austria (Mp9) is associated not only with its high landslide susceptibility but also with the underlying systematic mapping campaign associated with a recent project (Petschko et al., 2016). Between all the nine sub-inventories, the mapping domain named Mp1 covers the highest portion of the Austrian territory (73% of the area).

#### Geo-environmental variables

(Schweigl & Hervás, 2009) reported that landslide occurrence in Austria is mainly controlled by geo-environment factors such as

**Table 1** Characterization of available landslide information. The positional accuracy indicator (one to three crosses) was deduced from the authority's inventory data documentation. The positional accuracy was estimated higher (+++) in the case of inventories mapped utilizing high-resolution aerial photography, high-resolution DTM derivatives and GPS field campaigns. Mid accuracy (++) was assigned to digitalized inventories from other sources. Lower accuracy (+) was assigned to landslides inventories build on inferior accuracy mapping procedures (e.g. digitalization from old archives)

| Mapping domain (Mp) | Number of landslides and percentage (%) | Area (km$^2$) and national coverage percentage (%) | Landslide density (slide/km$^2$) | Mapping method technique and observations | Location describer in Austria | Qualitative positional accuracy indicator (accuracy suggested within the source documentation) |
|---|---|---|---|---|---|---|
| Mp1 | 8297 (34.73%) | 61,771 km$^2$ (73.5%) | 0.13 | Varying mapping techniques involved. Mostly digitalized historical inventories. Source: Geological Survey of Austria. | It covers the largest portion of the territory. It comprehends the provinces of Tyrol, Carinthia and Burgenland; and parts of Styria and Upper Austria. | + (1: 300,000) |
| Mp2 | 1212 (5.07%) | 2668 km$^2$ (3.18%) | 0.45 | Interpretation of multi-temporal aerial photography (Tilch, 2014). | Covers the Vorarlberg province area. | +++ (1: 2000) |
| Mp3 | 236 (0.99%) | 105 km$^2$ (0.13%) | 2.25 | Event-based inventory. Field mapping with GPS. Source: Geological Survey of Austria. | It is located nearby Feldbach municipality, province of Styria. | +++ (accuracy of a few meters) |
| Mp4 | 58 (0.24%) | 7 km$^2$ (0.01%) | 8.27 | Event-based inventory. Field mapping with GPS. Source: Geological Survey of Austria. | It is located nearby Feldbach municipality, province of Styria. | +++ (accuracy of a few meters) |
| Mp5 | 179 (0.75%) | 9 km$^2$ (0.01%) | 20.23 | Event-based inventory, digitalized after a field mapping performed by Brandlmayr (1995). | It is located in the municipalities of Gmunden and Ebensee, province of Upper Austria. | ++ (1: 50,000) |
| Mp6 | 21 (0.09%) | 115 km$^2$ (0.14%) | 0.18 | Event-based and site-specific mapping survey. Source: Geological Survey of Austria. | Located in the Hausruck region, province of Upper Austria. | +++ (1: 1000 to 1: 5000) |
| Mp7 | 390 (1.63%) | 60 km$^2$ (0.07%) | 6.49 | Event-based landslide inventory. Field mapping with GPS. Source: Geological Survey of Austria. | The landslides were mapped after a rainfall event in 2005 over the municipalities of Gasen und Haslau, both located in the Styria province. | +++ (>1: 10,000) |
| Mp8 | 443 (1.58%) | 12 km$^2$ (0.01%) | 36.52 | Event-based landslide inventory. | Event-based inventory mapped after the 2009 rainfall event | +++ (>1: 1000) |

**Table 1** (continued)

| Mapping domain (Mp) | Number of landslides and percentage (%) | Area (km²) and national coverage percentage (%) | Landslide density (slide/-km²) | Mapping method technique and observations | Location describer in Austria | Qualitative positional accuracy indicator (accuracy suggested within the source documentation) |
|---|---|---|---|---|---|---|
| | | | | Mapped through airborne laser scanning survey. Source: Geological Survey of Austria. | within the Klingfurth municipality, province of Lower Austria. | |
| Mp9 | 13,055 (54.64%) | 19,243 km² (22.91%) | 0.68 | Mapped through airborne laser scanning survey campaign (Petschko et al., 2014). | It comprehends mostly the whole extension of the Lower Austria province. | +++ (mapping based on a 1 m ALS-DTM) |
| Compiled inventory | 23,891 landslides (100%) | 83,991 km² (100%) | 0.28 | N.A. | N.A. | N.A. |

geological, geomorphological and land cover features. For this study, five topographical variables were derived from a resampled airborne laser scanning (ALS)–based DTM available at a 10 m × 10 m spatial resolution. Slope angle is the most commonly used variable in landslide susceptibility modelling (Budimir et al., 2015; Coe et al., 2004; Corominas et al., 2014; Kanungo et al., 2009; Malamud et al., 2014; Pourghasemi & Rossi, 2016; Süzen & Kaya, 2012; van Westen et al., 2008; Wu et al., 2015). This variable was selected to describe the driving forces that directly influence the downslope sliding potential. For large areas, elevation may reflect altitude-dependent environmental, climatic and morphological conditions that are associated with slope stability (e.g. general climatic features, elevation-dependent weathering variations) (Corominas et al., 2014; Dai & Lee, 2002; van Westen et al., 2008). The topographic wetness index (TWI) (Beven & Kirkby, 1979) was used as a hydrological proxy which considers the up-slope contributing area and slope gradient. Applied as a conditioning variable for this study, the TWI indicates potential spatial differences in soil moisture and flow accumulation (Süzen & Kaya, 2012).

The slope aspect may describe a varying degree of insolation, which influences soil moisture and weathering conditions (Corominas et al., 2014; Dai & Lee, 2002; van Westen et al., 2008). The slope aspect is also mentioned to be a relevant land-slide predisposing variable considering that distinctive aspects may influence the terrain exposition to rainfalls and radiation, conditioning terrain humidity and vegetation patterns (Catani et al., 2013; Pourghasemi & Rossi, 2016). For this study, the aspect was derived from the DTM and resampled from the continuously scaled layer (originally from 0 to 360°) by calculating the respective cosine and sine, representing the degree of north and east exposedness (Brenning, 2009; Brenning & Trombotto, 2006; Steger et al., 2016b). The lithological map provided by the Geological

Survey of Austria at the scale 1:500,000 (Weber, 1997) was used as an indicator for the parent structural material composition. The lithological classes were represented by their lithostratigraphical units. The land cover data, used to account for an associated inventory-based incompleteness (cf. "Statistical modelling"), was obtained from the CORINE land cover data (CLC) (European Environment Agency & EEA, 2012).

**Methods**

The methodological framework (Fig. 3) consists of four main steps: (i) data collection and preparation, (ii) exploratory data analysis, (iii) statistical modelling and (iv) model evaluation. Four models were created by testing two classifiers (logistic regression, mixed-effect logistic regression) and two landscape representations (grid-based, slope unit–based). The slope units were semi-automatically delimited using GRASS GIS and *r.slopeunits* (Alvioli et al., 2016). All statistical analyses were performed using the open-source statistical software "R" (Core Team, 2020). The terrain and thematic GIS parameters were computed using the open-source "SAGA GIS" (Conrad et al., 2015), while final visualizations of the maps were conducted within ESRI ArcGIS (ArcGIS Desktop, 2017).

**Data collection and preparation**

Two different landscape representations were tested. For the grid-based approach, all the variables were resampled to the modelling resolution of 100 m × 100 m. The delineation of slope units was performed using the *r.slopeunits* extension of GRASS GIS (Alvioli et al., 2016) using a scale-dependent parameter optimization, as suggested by (Schlögel et al., 2018). The geo-environmental variables were assigned to the underlying slope units according the following criteria: for continuous variables (e.g. slope and elevation), the mean value was taken; while for categorical variables
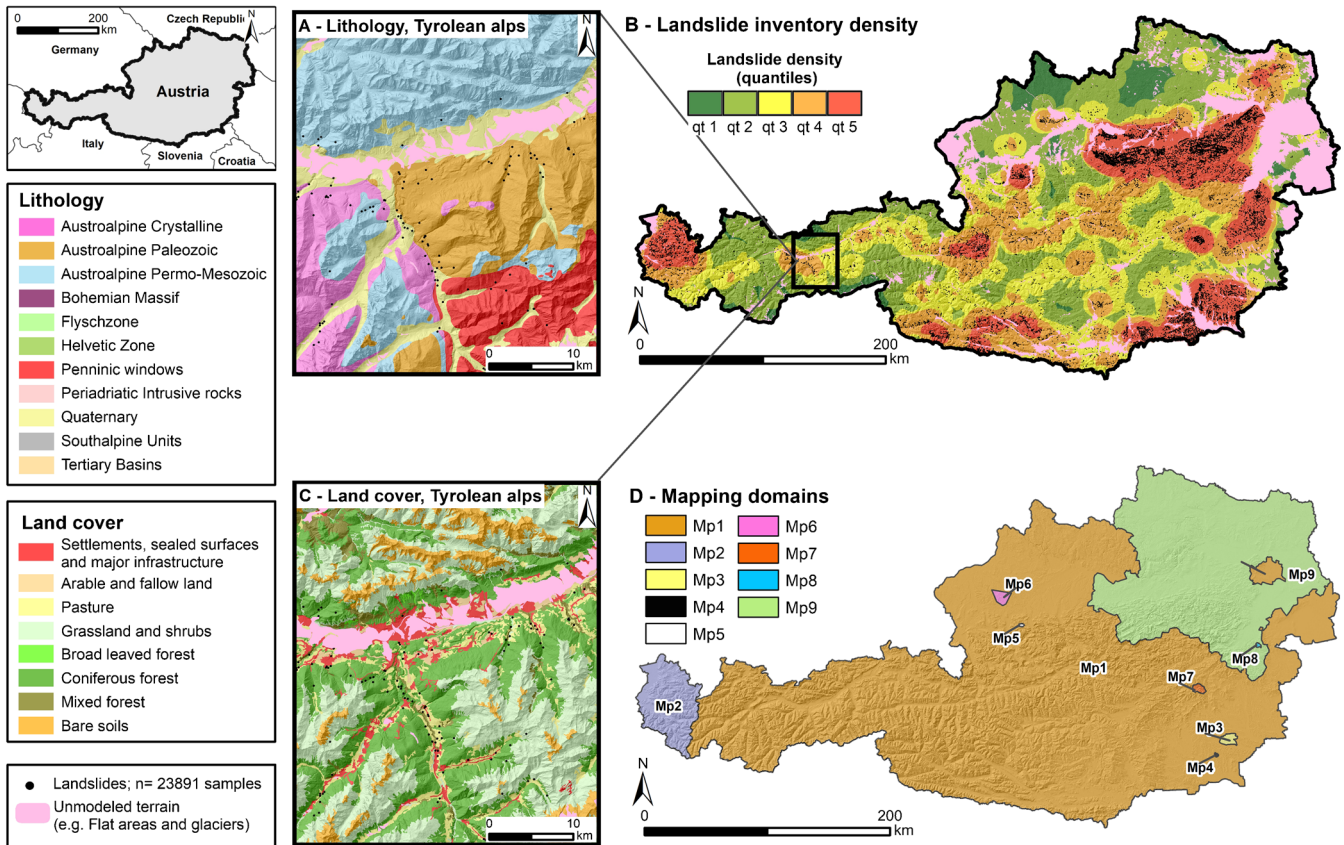
**Fig. 2** The landslide inventory composed of 23,891 observations is displayed within the central image as black points which are the basis for the calculated landslide density (B). Two closer frames display details of the lithology (A) and land cover variables (B), respectively. The nine different inventory mapping domains (Mp) are given in (D)

(e.g. land cover and lithology), the predominant class was taken. Slope units that were entirely located within the class "flat", as depicted by the topographical position index (Weiss, 2001), or were predominantly located within fluvial deposits, as depicted by the lithological map, were not considered for modelling. The exclusion of such flat "trivial terrain" was expected to increase the explanatory power of the results at the costs of an apparent lower predictive performance (Steger & Glade, 2017). These trivial areas were also excluded for the grid-based analyses. To guarantee a parsimonious model and to enhance the interpretability of the results, we opted to build the study upon a compact set of frequently applied geo-environmental variables (cf. "Geo-environmental variables").

### Exploratory data analysis

An initial examination of the available datasets built the basis to gain insights into empirical relations between geo-environmental variables and the presence/absence of landslide occurrence. The data visualization techniques were also used to explore suspicious data patterns, which may be indicative of data biases. Conditional frequency plots (for continuous variables) and spineplots (for categorical variables) were used to highlight the ratio of landslide presence to absence across the geo-environmental data values. First insights into the capability of geo-environmental variables to distinguish landslide

presence from absence observations were gained by evaluating the discriminatory power of single-variable models (Murillo-García et al., 2019; Steger et al., 2020). In this case, the obtained metric reflects the fitting performance of a single-predictor logistic regression measured via the area under the receiver operating characteristic curve (AUROC) (Beguería, 2006; Remondo et al., 2003).

### Statistical modelling

Regression-based classification methods are frequently applied to model landslide susceptibility (Brenning, 2005; Budimir et al., 2015; Malamud et al., 2014; Wu et al., 2015). Two classifiers, which are both based on a generalized linear model (GLM), were confronted for this study: logistic regression and mixed-effects logistic regression. The former relates to the most widely applied classifier in landslide susceptibility modelling, while the latter is rather new in the field and allows to additionally isolate variation related to a heterogeneous landslide inventory completeness among classes of categorical variables (i.e. mapping domains, land cover units) (Steger et al., 2017).

Atkinson and Massari (1998) and Guzzetti et al. (1999) were among the first to apply logistic regression (LR) for spatial landslide susceptibility modelling. Besides producing smooth prediction surfaces, the LR results are straightforward to interpret (Felicísimo et al., 2013; Goetz et al., 2015). The applied LR model

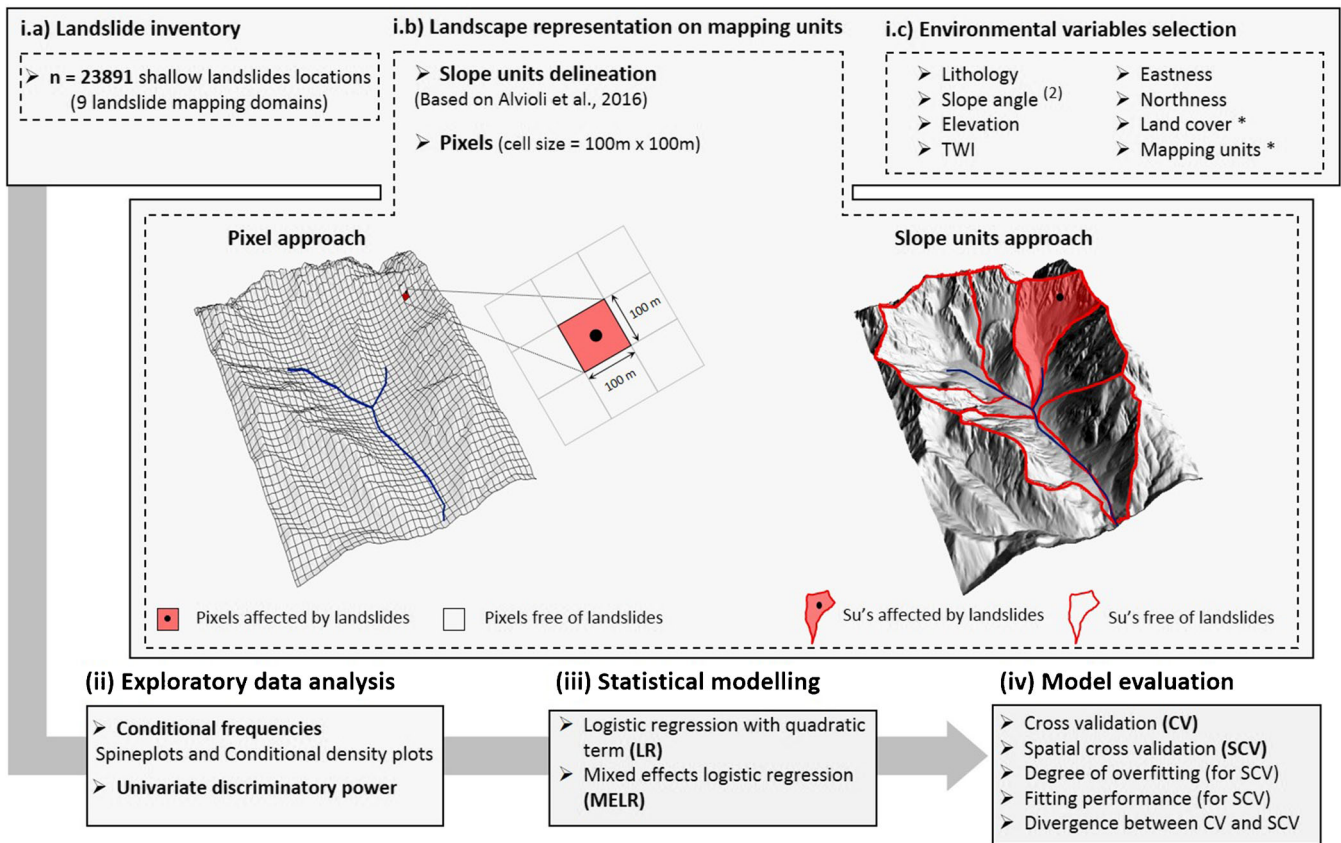## (i) Data collection and preparation

**i.a) Landslide inventory**

> **n = 23891** shallow landslides locations
> (9 landslide mapping domains)

**i.b) Landscape representation on mapping units**

> **Slope units delineation**
> (Based on Alvioli et al., 2016)

> **Pixels** (cell size = 100m x 100m)

**i.c) Environmental variables selection**

> Lithology
> Slope angle [(2)]
> Elevation
> TWI

> Eastness
> Northness
> Land cover *
> Mapping units *



**Pixel approach**

100 m
100 m

■ Pixels affected by landslides    □ Pixels free of landslides

**Slope units approach**

◆ Su's affected by landslides    ◇ Su's free of landslides

**(ii) Exploratory data analysis**

> **Conditional frequencies**
> Spineplots and Conditional density plots
> **Univariate discriminatory power**

**(iii) Statistical modelling**

> Logistic regression with quadratic term (**LR**)
> Mixed effects logistic regression (**MELR**)

**(iv) Model evaluation**

> Cross validation (**CV**)
> Spatial cross validation (**SCV**)
> Degree of overfitting (for SCV)
> Fitting performance (for SCV)
> Divergence between CV and SCV

**Fig. 3** Methodological framework. (i) Data collection and preparation. (i.a) National-scale landslide inventory assemble. (i.b) Landscape representation through terrain units (grid approaches and slope units). (i.c) selection, preparation and representation of the variables through the selected landscape representation. (ii) First knowledge gain on the data through the exploratory data analysis. (iii) Statistical classifiers applied. (iv) Outcome evaluation and best performing model selection

assumes a binomial probability distribution of the response and the final output relates to the probability of landsliding (Hosmer & Lemeshow, 2000). LR was fitted using six predictors: lithology, slope angle, elevation, TWI, eastness and northness. To better represent the previously explored non-linear relationship between shallow landslide occurrence and slope angle (cf. "Exploratory data analysis"), a quadratic term ($X^2$) was applied (Osborne, 2015).

Most statistical classifiers used in the field of landslide susceptibility modelling, such as LR, can be assigned to the category of fixed-effects models which aim to assess the direct influence of each predictor variable (i.e. the fixed-effect) on the response (Bolker et al., 2009; Steger et al., 2017). Mixed-effects models enable to additionally consider random effects in order to account for data hierarchies or nuisance effects (Bolker et al., 2009; Zuur et al., 2009). For statistical landslide susceptibility modelling, mixed-effects modelling already proved efficient to separate effects related to a systematic incompleteness inherent in the landslide data from effects that describe the quantity of interest, namely landslide susceptibility (Steger et al., 2017). The random intercept was only used for parameter estimation and averaged-out for the final spatial prediction (Bolker et al., 2009; Zuur et al., 2009). Table 2 summarizes the selected variables applied within the LR and MELR models. It is important to note that the random intercepts (i.e. bias-describing variables) were only used for parameter

estimation, and lately averaged-out of the final predictions. The final spatial predictions are based on the fixed-effects variables (i.e. slope angle, lithology, elevation, TWI and aspect; eastness and northness). More details on mixed-effect modelling for assessing landslide susceptibility can be found in (Steger et al., 2017). Rather than simply ignoring biased variables, the advantage of mixed-effects model adoption is closely related to the confounding

**Table 2** Summary of the variables considered within this publication to fit and predict the statistical models

| Variables | Usage in the model | |
|---|---|---|
| | LR | MELR |
| Lithology | Fit and predict. | Fit and predict. |
| Slope angle | Fit and predict. | Fit and predict. |
| Elevation | Fit and predict. | Fit and predict. |
| TWI | Fit and predict. | Fit and predict. |
| Eastness | Fit and predict. | Fit and predict. |
| Northness | Fit and predict. | Fit and predict. |
| Land cover | Fit and predict. | Fit and random intercept. |
| Mapping units | Not used. | Fit and random intercept. |

effects, often observed between environmental variables (e.g. slope and land cover). Therefore, the usage of bias-describing variables as random intercepts enables to account for the associated variation during model parameter estimation (Steger et al., 2021), while the effects of these same variables are averaged-out for the predictions. For this study, the two categorical variables, land cover and mapping domain, which mainly describe a systematic incompleteness of landslide information, were introduced as random intercepts to reduce associated confounding and direct bias propagation.

Land cover variables are well known to describe not only landslide influencing processes but often also a heterogeneous completeness of landslide information (Bell et al., 2012; Petschko et al., 2016). For instance, forest cover may impede the identification of landslides while an over reporting of landslide information is likely nearby settlements or agricultural land (Bell et al., 2012; Brardinoni et al., 2003; Petschko et al., 2016; Steger et al., 2017). The completeness of landslide information was also expected to vary across the nine mapping domains (cf. "Landslide inventory"). Thus, the nine mapping domains (Table 1; Fig. 2(D)) were also introduced as a random intercept within the MELR.

### Model evaluation

Model evaluation is considered an essential step in landslide susceptibility modelling (Chung & Fabbri, 2003; Guzzetti et al., 2006). The AUROC was used as a performance metric and assessed for the training samples (fitting performance) and test samples (predictive performance) obtained by repeated non-spatial (cross-validation (CV)) and spatial partitions (spatial cross-validation (SCV)). The training and test sample partitioning was performed using a k-Fold cross-validation technique using 25 repetitions and 10-fold for each repetition. This multi-fold partitioning technique is described in more detail within (Brenning, 2012; Schratz et al., 2019). In summary, for each model 250 AUROC values that relate to different partitions of training and test data were calculated (25 repetitions times 10 folds). Model overfitting describes the tendency of a model to adapt itself too closely to the training sample and therefore fails to explain independent test set observations (Hosmer & Lemeshow, 2000). An overfitted landslide susceptibility model may reproduce the training observations in great detail, while yet unseen future landslide locations may remain undetected (Brenning, 2005; Goetz et al., 2015). The difference between the fitting and predictive performances of each model was calculated to get insights into the index of model overfitting. (Steger et al., 2017) highlighted that considerable differences between non-spatially (CV) and spatially (SCV) assessed predictive performances can indicate systematic spatial inconsistencies in the modelling results. Thus, mean AUROC differences ($\Delta|CV - SCV|$) were calculated to expose inconsistent modelling results. Besides a detailed quantitative model evaluation, also a geomorphological plausibility check was conducted to explore whether the results were affected by evident inventory-based biases or artefacts (Steger et al., 2016a).

The geomorphological plausibility of landslide susceptibility models might suffer in case the model training is based on biased landslide data, such as underreporting of past events within specific territories. Taking the landslide data background into account, subsequent bias propagations and misleading spatial predictions can be identified. The morphological coherence of the maps was assessed qualitatively by considering known and suspected flaws in the available landslide data and by scrutinizing whether those flaws are reflected in the final prediction pattern. As an example, a section in the Tyrolean Alps (Fig. 2(A, D); Fig. 6) was selected to check the coherence of the results. This landslide-prone area is characterized by a high relief energy (540–3166 m asl.; mean elevation of 1638 m asl.) and steep terrain (mean slope 25°, max. 76°) and known to be underrepresented in terms of available landslide information. Predicted very low landslide susceptibility scores at the prevalent hill slopes were therefore interpreted as an indicator of a direct landslide data bias propagation and a low morphological coherence of the final map.

## Results

### Exploratory data analysis

The exploratory analysis (Fig. 4) shows comparable conditional frequencies for the grid-based and slope unit–based terrain representations. For both landscape representations, the plots for the variable lithology evidence the highest conditional landslide frequencies for the units Flysch (Fz), Helvetic zone (Hz) and South Alpine (Sa). Comparably low landslide densities were observed for the units Penninic window (Pw) and Bohemian Massif (Bm). High conditional landslide frequencies were calculated for the land cover class pastures (P), broad-leaved forests (Bf) and mixed forests (Mf). Bare soils (Bs) showed low landslide frequencies. The highest landslide frequencies were observed for medium inclined terrain, with the highest values between 10 and 30° (Fig. 4(D, H)).

The land cover plots (Fig. 4(B, F)) also reveal a comparably low number of inventoried landslides for settlement areas and arable land. Since these land cover units are simultaneously associated with lower slope angles and low elevations, such univariate data inspections have to be interpreted with care. An additional consideration of the underlying landslide data origin (e.g. mapping purpose) indicates that the observed relations (e.g. land cover vs. landslide occurrence) do not necessarily depict geomorphically plausible relations but are likely to describe a spatially heterogeneous landslide data completeness. Higher landslide frequency was also associated with low to mid slope steepness and lower to medium elevation (Fig. 4), where usually most of the settled areas tend to be located. This supports the argument of an overrepresentation of landslides near settlements. Lower frequencies of landslides were associated with land cover features like arable land, grassland, coniferous forest and bare soils. Predominantly located at steeper slopes, coniferous forests were associated with the lowest frequency of reported landslides compared to the other forest types.

The discriminatory power of each variable varied from 0.5 (TWI and Eastness when applied for the grid approach) to 0.73 (for lithology when applied for the grid approach, refer to Table 3). For the grid-based approach, lithology and slope angles were most efficient in discriminating landslide presence from absence. The discriminatory power was generally lower when assessed for the slope unit models.

### Spatial prediction maps

It was observed that the appearance of the predictive maps differed substantially. In general, the MELR classifier
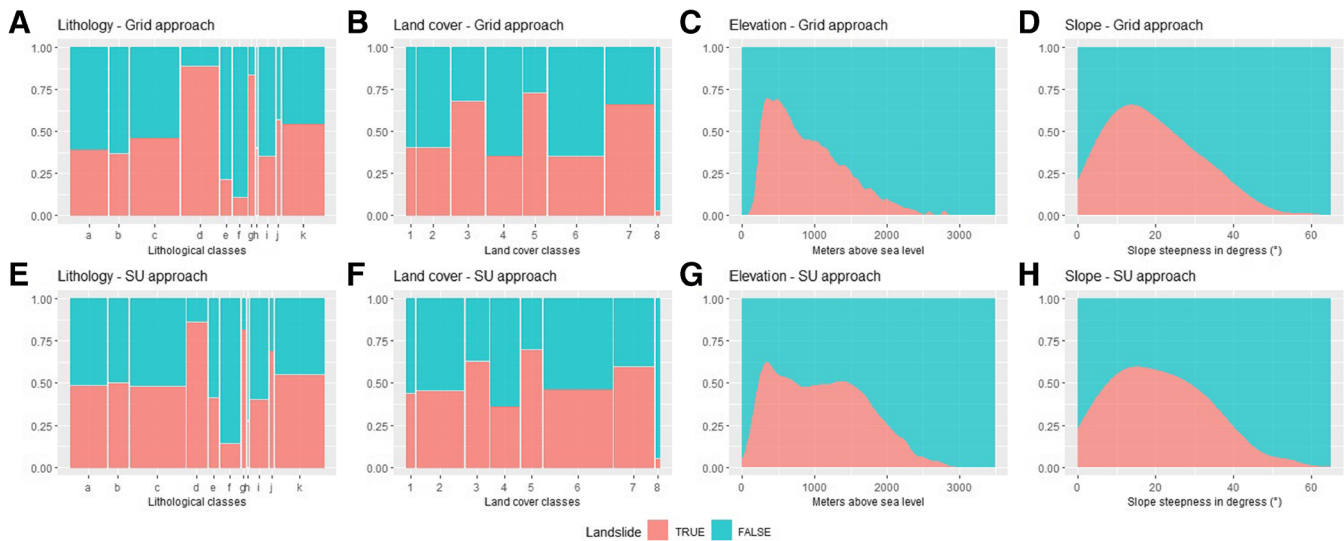
**Fig. 4** Conditional frequency plots. Graphs (A)–(D) represent the grid approach, and graphs (E)–(H) represent the slope unit approach. The light orange colour represents landslide presence (TRUE), while the light-blue colour represents the non-landslide observations (FALSE). The y-axis represents either the binary or the relative proportion between landslide and non-landslide. Legend for the internal variables categories: Graphs (A) and (E) - Lithology: (a) Austroalpine crystalline, (b) Austroalpine Paleozoic, (c) Austroalpine Permo-Mesozoic, (d) Flysch zone, (e) Penninic window, (f) Bohemian Massif, (g) Helvetic window, (h) Periadriatic rocks, (i) Quaternary, (j) South Alpine units, (k) Tertiary basis. Graphs (B) and (F) - Land cover: (1) = settlements, (2) = arable and fallow land, (3) = pasture, (4) = grassland, (5) = broad-leaved forests, (6) = coniferous forests, (7) = mixed forests, (8) = bare soils. For the spine plots (graphs (A), (B), (E) and (D)), the width of the bars corresponds to the relative spatial coverage of each variable class within the Austrian territory

produced a less spatially contrasting prediction pattern over the territory. LR predicted clear heterogeneous landslide susceptibility maps. The LR models clearly reproduced the higher and lower landslide distribution pattern observed in Fig. 2(B).

Although this general pattern can also be observed when applying MELR on slope units (Fig. 5), it is much more accentuated when potential bias-causing effects are ignored (i.e. LR models).
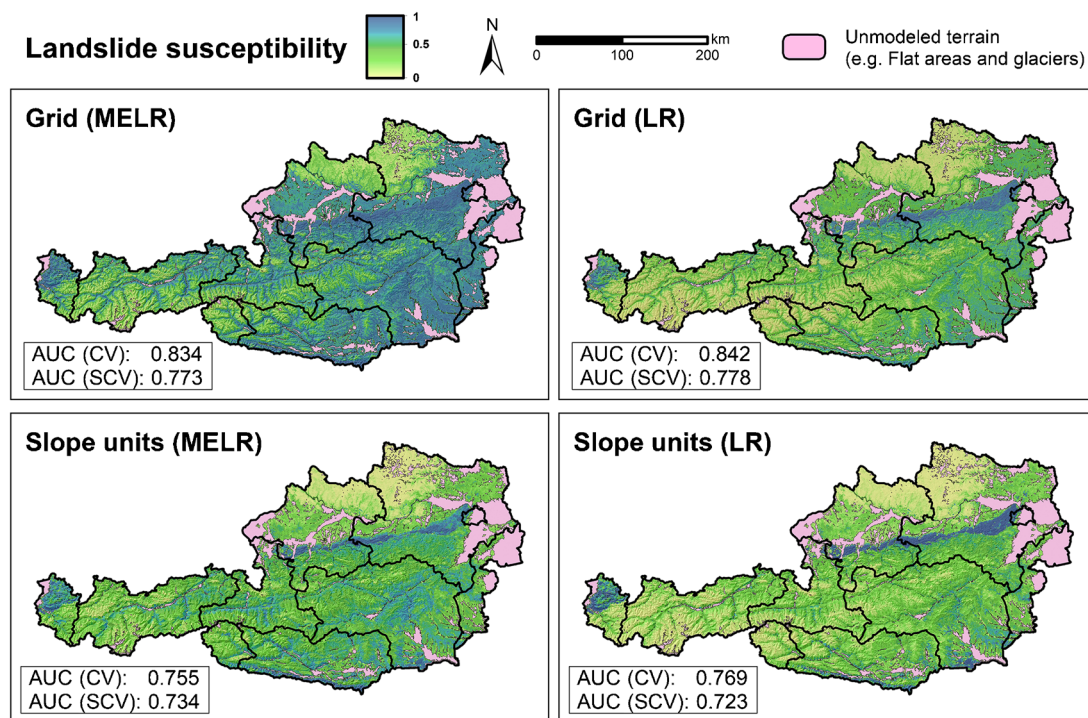


**Fig. 5** National-scale landslide susceptibility maps. The four models were produced under diverse statistical classifiers (LR or MELR) for grid and slope unit approaches. The predictive performance scores (median AUROC) are represented on the bottom-left corner. The upper maps represent the grid approaches (predicted through MELR and LR), while the bottom maps represent the slope unit models (predicted through MELR and LR). In order to avoid any additional bias, some areas have not been modelled (e.g. flat areas, glaciers)

**Table 3** Univariate AUROCs associated with each single-predictor model

| Variables | Modelling approaches | |
|---|---|---|
| | Grid approach | Slope units |
| Lithology | 0.73 | 0.66 |
| Land cover | 0.68 | 0.62 |
| Slope | 0.67 | 0.64 |
| Elevation | 0.68 | 0.58 |
| Northness | 0.53 | 0.53 |
| Eastness | 0.51 | 0.51 |
| TWI | 0.54 | 0.54 |

A closer look at the high alpine areas of the Tyrolean Alps indicates considerable discrepancies in the produced spatial prediction patterns when comparing the different modelling strategies (Fig. 6). In contrast to MELR, LR models tend to produce low landslide susceptibility scores even for steeper parts of the Tyrolean Alps. This behaviour became particularly evident when comparing landslide susceptibility scores across slope angles for this same mountainous region (Fig. 8). Instead, for this same region and despite the low amount of inventoried landslides in this area, MELR was still able to assign comparably high susceptibility scores
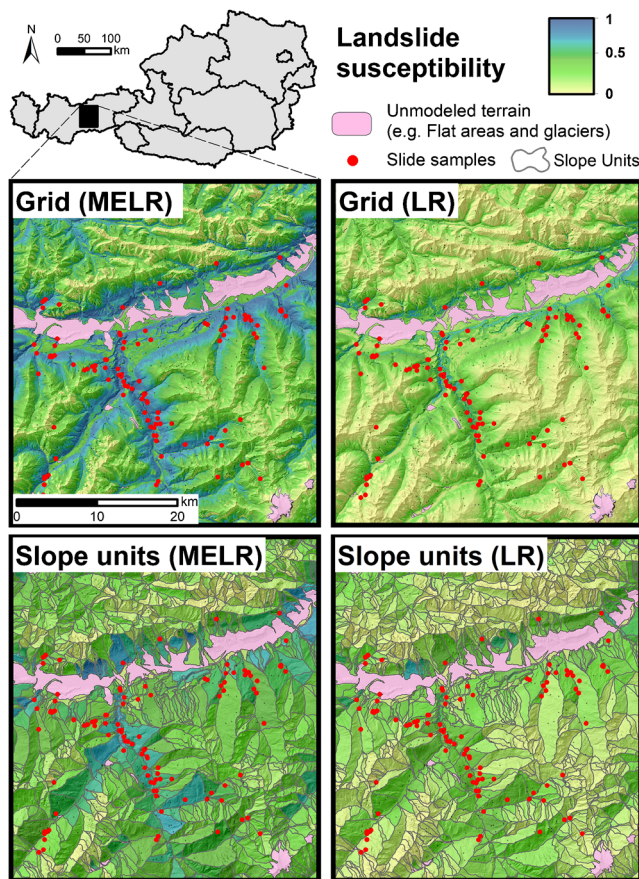


**Fig. 6** Closer cutout to a region with a low density of landslides. The upper maps represent the grid approach (predicted through MELR and LR, respectively), and the bottom maps represent the slope unit approach (also respectively predicted through MELR and LR).

to the steep terrain of this particular area. For this area, 94% of the pixels were classified lower than 0.5 when applying LR. When accounting for potential inventory bias using MELR, the percentage of pixels predicted lower than 0.5 decreased to 53%.

**Model evaluation**

Among the grid-based approaches, LR reached the highest predictive performance (CV: 0.842; SCV: 0.778) compared to MELR (CV: 0.834; SCV: 0.773). Also, for the slope unit models based on LR, the AUROCs (CV: 0.769; SCV: 0.723) were slightly higher than their MELR equivalents (CV: 0.755; SCV: 0.734). Lower predictive performances were constantly obtained for the spatial cross-validation technique (SCV) compared to non-spatial cross-validation (CV) as also observed by Petschko et al., (2014), Steger & Glade, (2017), Steger et al., (2016b).

The fitting performance achieved through SCV (black crosses in Fig. 7) showed a similar trend with slightly higher values for the LR models. Fitting performances can jointly be interpreted with the predictive performance to obtain insights into the index of model overfit (here named as overfitting index and represented by the black points within the lower graph of Fig. 7) (Brenning, 2005; Goetz et al., 2015; Murillo-García et al., 2019; Tien Bui et al., 2012). Comparing the two classifiers, the overfitting index was lower for MELR indicating a lower index of overfitting. This tendency was particularly evident for the slope unit terrain partition. The mean difference between CV and SCV ($\Delta|CV - SCV|$), in Fig. 7, was constantly lower for MELR indicating spatially more robust modelling results (Steger et al., 2017).

**Discussion**

This research tackled two pending challenges related to the topic of large area statistical landslide susceptibility assessment, namely (i) the systematic incompleteness of landslide data and (ii) the unprecise positional location of landslide samples. These challenges were faced using (i) a mixed-effects modelling approach and (ii) an alternative representation of the terrain, namely slope units.

The initial exploratory data analysis provided further evidence that systematic biases are inherent in the available landslide data. The inventory data that was based on reports was assessed to overrepresent densely populated areas (gentle slopes and relatively lower elevations), which is reflected by a very high conditional frequency of landslides (Fig. 4 (C, D, G, H)). As a consequence, an underreporting of landslides in sparsely populated regions can be expected. Particular high conditional landslide frequencies for both terrain units were associated with elevations below 1500 m and slope inclinations below 20 degrees. Additionally, known challenges in landslide reporting and mapping within specific land cover classes (e.g. forested areas) as described by (Bell, 2007); (Petschko et al., 2016); (Conoscenti et al., 2016) might have contributed to a heterogeneous landslide data completeness. This in turn supported the decision to include land cover as a bias-describing effect within the mixed-effects models.

While the four models reached moderately high and similar predictive performances, the appearance of the final maps varied significantly. Such contrasting spatial prediction patterns for similarly performing models have also been reported in the literature (Hussin et al., 2016; Sterlacchini et al., 2011). The slightly higher predictive performance of the LR models should be interpreted in
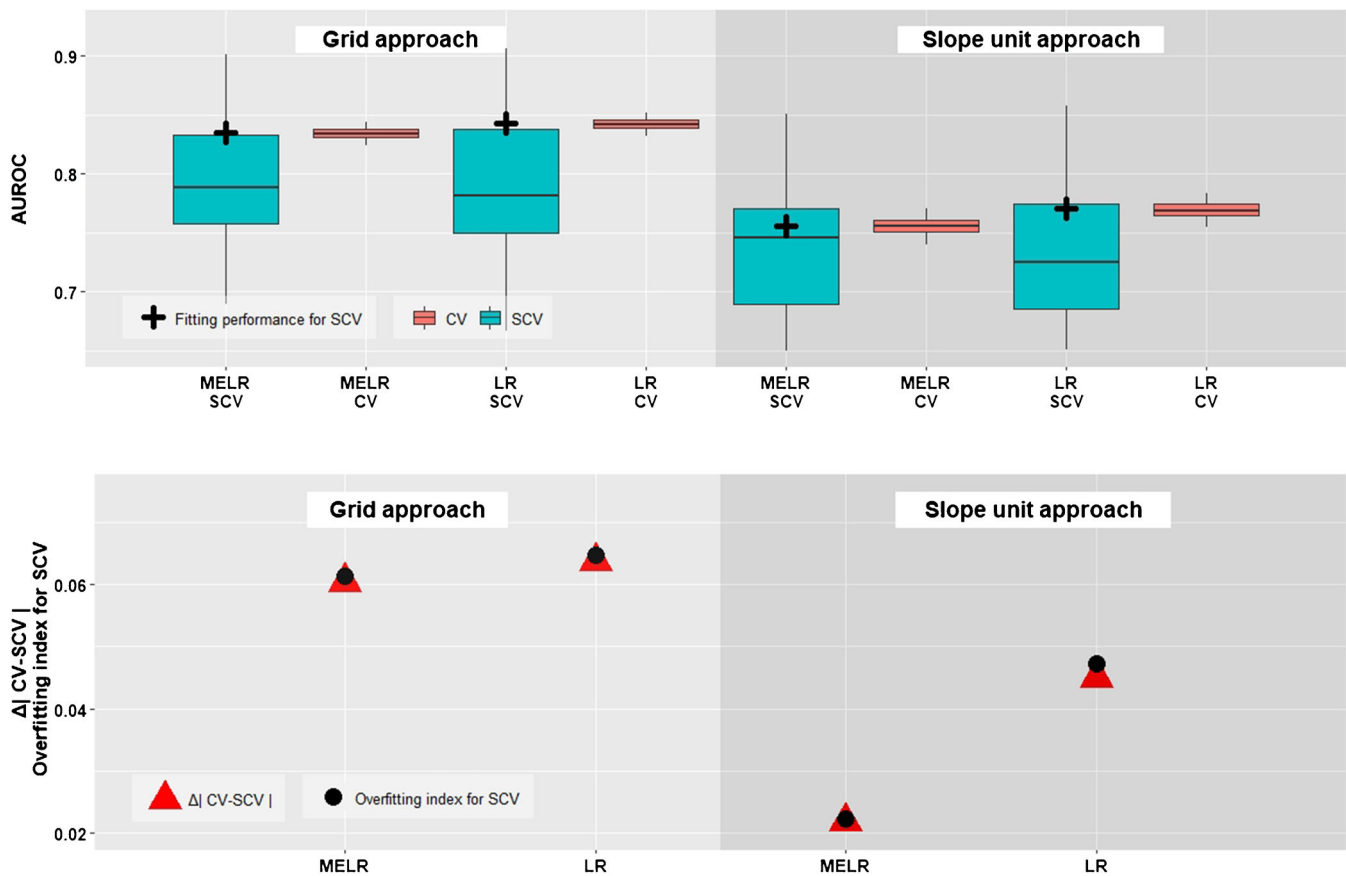
**Fig. 7** Quantitative model evaluation based on the spatial (SCV) and non-spatial cross-validation (CV). The boxplots in the upper graphs represent the CV- and SCV-based AUROCs, while the black cross symbol refers to the median SCV fitting performance. The graph at the bottom represents the overfitting index calculated for SCV partitions (black points) and an indicator for spatially inconsistent modelling results, i.e. the difference between median spatial and non-spatial validation scores (red triangles)

the context of the underlying landslide data bias. By including a bias-describing predictor like land cover, LR directly reproduced the associated data bias which also led to over-optimistic AUC values (Goetz et al., 2015; Hosmer & Lemeshow, 2000). Although a vast amount of publications focuses on the AUC values as a model and variable selection criterion, high AUCs might not necessarily reflect the real quality of the final maps, especially under data bias conditions (Steger et al., 2016b; Steger et al., 2017). From a purely quantitative point of view, MELR models performed slightly worse compared to LR, a classifier extensively used in the field (Akgun, 2012; Atkinson & Massari, 2011; van Den Eeckhaut et al., 2012; Goetz et al., 2015; Guzzetti et al., 1999; Lee et al., 2004; Moosavi & Niazi, 2016; Nefeslioglu et al., 2008; Pourghasemi et al., 2013; Regmi et al., 2014; Reichenbach et al., 2014; Reichenbach et al., 2018; Trigila et al., 2013). Steger et al. (2017) have shown the potential of mixed-effects modelling for handling landslide data bias for study sites in Austria. This study provides evidence that the application of such approaches is beneficial, also when applied for large area assessments (national scale). MELR achieved was able to reduce the propagation of biased relationships into the final results and produced geomorphologically coherent predictions.

Within this research, the quality of the models was not only interpreted from calculated predictive performance estimates, but also on the basis of other indicators, such as the overfitting index and the difference between CV and SCV. Fitting a model too

closely to characteristics in training data is a common concern for landslide susceptibility models using statistical methods (Goetz et al., 2015). Such model overfitting was constantly observed to be lower for the MELR models (Fig. 7) compared to their LR counterparts. At the same time, MELR models were also associated with a lower difference in AUCs between the validation techniques (e.g. $\Delta|CV - SCV|$; Fig. 7), indicating a higher spatial consistency of the results (Steger et al., 2017).

Moreover, beyond the numerical evaluation, the model selection was additionally conducted by assessing the geomorphic plausibility of the maps (Bell, 2007; Steger et al., 2016a). This might be particularly important because differently appearing maps can be associated with similar performance estimates and the appearance of the final maps co-determines the practical acceptance by the final users (Brenning, 2005; Goetz et al., 2015). A close inspection of the prediction patterns created for the Austrian territory (Fig. 5) provided valuable evidence that the inventory biases were directly propagated into the final LR results. In fact, the lack of spatially consistent landslide data across the sampled target area (Fig. 2) is a recurrent challenge in the field of landslide susceptibility assessments, especially for very large areas. The inclusion of landslide mapping domains as a random effect within MELR counterbalanced the associated data heterogeneity across the territory and enabled more plausible results.

In order to assess how landslide data incompleteness affected the final maps, we focused on regions where the landslide occurrence is well known to be underestimated due to underreporting. For the selected landslide-prone Tyrolean area, only a few landslides were registered within the available inventory (samples originally from Mp1), which is mostly reported biased. When viewing the maps from a national-scale perspective (Fig. 5), it became obvious that LR reproduced this data bias by assigning particularly low susceptibility scores to this area (Fig. 6). From a geomorphic viewpoint, much higher landslide susceptibility scores can be expected for this particular region. The approach based on mixed-effects modelling was able to counterbalance this landslide data flaw and produced more appropriate predictions (i.e. higher susceptibility scores) for the hillside areas. In other words, even in the case that very few landslides were officially registered for the Tyrolean hill slope areas, MELR assigned relatively high probability scores as a consequence of the prevalent morphology.

For this sub-region, the lower plausibility of the LR predictions became particularly evident when plotting the predicted susceptibility values against a factor directly related to shear stress, namely the slope angle (Fig. 8). A more detailed view at the prediction patterns for this region (Fig. 8(B)) shows for LR a particularly high concentration of very low susceptibility scores (close to zero) associated even with high slope inclinations (between 30 and 40 degrees). MELR in its place produced a more balanced representation of landslide susceptibility for this region by predicting a substantial amount of medium to steeply inclined slopes as considerably susceptible to landsliding (Fig. 8(B)).

The application of slope units as an alternative to grid-based assessments has recently gained more and more attention in the field of landslide susceptibility assessment (Alvioli et al., 2016; Camilo et al., 2017; Guzzetti & Reichenbach, 1994; Jacobs et al., 2020; Lombardo et al., 2018; Reichenbach et al., 2014; Schlögel et al., 2018). Our results are in line with (van Den Eeckhaut et al., 2009) and demonstrate an overall higher predictive

performance for grid-based assessments, in comparison to their slope unit counterparts. However, other studies showed that this is not necessarily always the case (Erener & Düzgün, 2012). A common aspect of all publications that compared both landscape representations (slope units vs. grid cells) for the purpose of landslide susceptibility modelling is the observed similar predictive performance of the underlying models, despite the rather different appearance of the final maps (e.g. van Den Eeckhaut et al. (2009)). This already suggests that focusing on obtained performance metrics as the sole criteria to select a model is subject to limitations.

For landslide susceptibility assessments, slope units may frequently cover a larger areal extent than a pixel within a conventional grid-based model. As a consequence, inaccurately mapped landslides are still more likely to be assigned to the correct spatial entity (i.e. slope unit) compared to their grid-based counterparts (Steger et al., 2016b). Thus, slope unit–based models are likely to be less sensitive to positional inaccuracies of inventory data. (Jacobs et al., 2020) tackled the effects of uncertain landslide point positioning on landslide susceptibility models and confirmed an improved capacity of slope units to handle positionally inaccurate landslide data, compared to pixel-based representations. Within this research, slope units, with generally larger size than pixels, were better able to accommodate positional inaccuracies from the inventories while still being a reliable terrain unit for landslide susceptibility models. Finally, the joint analysis of several evaluation criteria (prediction pattern surfaces, the susceptibility frequencies distribution and the different validations measures) provides further evidence of the utility of slope units under landslide data bias conditions and positional inaccurate sampling points.

Ultimately, the current results provide evidence that a too detailed representation of the terrain may be detrimental in the presence of inaccurate landslide data and that actively counterbalancing known systematic data biases (e.g. averaging out bias-describing variables using mixed-effects modelling) can
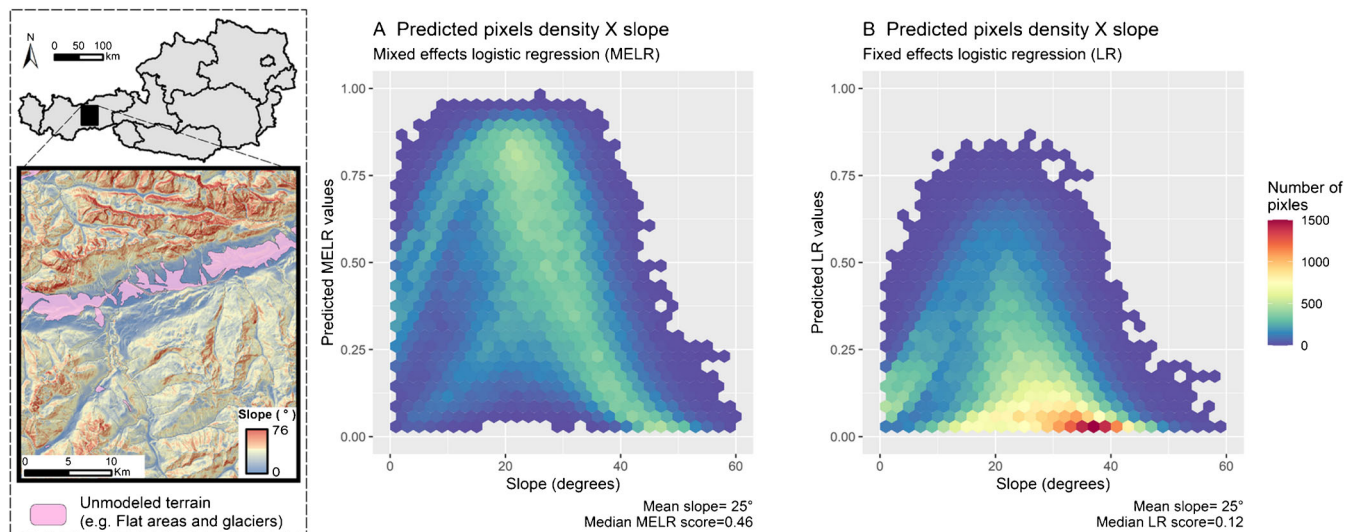


**Fig. 8** Comparison between predicted susceptibility scores (y-axis in (A) and (B)) and slope angle for MELR (left-side graph (A)) and LR (right-side graph (B)). The graphs and summary statistics at the bottom of each graph are related to the shown area

improve the plausibility of the results. In this context, we consider mixed-effects models in combination with slope units valuable to handle the impact of flawed landslide information.

## Conclusion

Landslide inventory data available for large areas is usually affected by positional inaccuracies and spatial incompleteness. For national-scale analyses, the common unavailability of accurate and representative information on past slope instabilities impedes the straightforward creation of statistically based landslide susceptibility models. This research highlighted that an adaptation of the research design can minimize the propagation of landslide data flaws into susceptibility models for very large areas. The underlying comparative analyses were based on four models which are related to different classifiers (conventional logistic regression vs. mixed-effects logistic regression) and different terrain representations (grid-based vs. slope unit–based).While conventional logistic regression did not specifically account for the underlying data bias, a mixed-effects modelling approach was applied to counterbalance effects associated with a systematic spatial landslide data incompleteness. Using slope units, instead of the more common pixel-based terrain representation, allowed to reduce the effects of positionally inaccurate landslide locations. A holistic evaluation of modelling results (i.e. quantitative and qualitative assessments) provided evidence that mixed-effects modelling in combination with a slope unit terrain representation was beneficial under the prevalent flawed landslide data conditions compared to the standard procedures (e.g. logistic regression and a grid-based terrain representation).

For large area landslide susceptibility assessment, we recommend to (i) gain insights into potential landslide data flaws in order to (ii) allow a corresponding adaptation of the modelling design. In case the landslide data is heterogeneously complete across an area, we advise to avoid explanatory variables that describe and therefore reproduce the underlying landslide data incompleteness. Instead, mixed-effects modelling can prove useful to explicitly reduce associated biases. Avoidance of a detailed representation of the terrain (e.g. via high-resolution grid-based models) is beneficial to tackle the challenge of positionally inaccurate landslide information. In this context, we advocate considering an alternative representation of the terrain, such as slope units. We finally emphasize that such large area analyses have to be interpreted with care, even if the flaws inherent in the data were accounted for in the research design. The present results provide a generalized overview of landslide-prone areas in Austria, but they are not applicable for local decision-making.

## References

Akgun A (2012) A comparison of landslide susceptibility maps produced by logistic regression, multi-criteria decision, and likelihood ratio methods: a case study at Izmir, Turkey. Landslides 9(1):93–106. https://doi.org/10.1007/s10346-011-0283-7

Alvioli M, Marchesini I, Reichenbach P, Rossi M, Ardizzone F, Fiorucci F, Guzzetti F (2016) Automatic delineation of geomorphological slope-units and their optimization for landslide susceptibility modelling. Geoscientific Model Development Discussions pp:1–33. https://doi.org/10.5194/gmd-2016-118

ArcGIS Desktop (2017) ArcGIS Release 10.5.1 Environmental Systems Research Institute, Redlands

Ardizzone F, Cardinali M, Carrara A, Guzzetti F, Reichenbach P (2002) Uncertainty and errors in landslide mapping and landslide hazard assessment. Natural Hazard and Earth System Science 2(1–2):3–14

Atkinson PM, Massari R (1998) Generalised linear model of susceptibility to landsliding in the central Apennines, Italy. Comput Geosci 24(4):373–385. https://doi.org/10.1016/S0098-3004(97)00117-9

Atkinson PM, Massari R (2011) Autologistic modelling of susceptibility to landsliding in the Central Apennines, Italy. Geomorphology 130(1):55–64. https://doi.org/10.1016/j.geomorph.2011.02.001

Auer I, Böhm R, Jurkovic A, Lipa W, Orlik A, Potzmann R, Schöner W, Ungersböck M, Matulla C, Briffa K, Jones P, Efthymiadis D, Brunetti M, Nanni T, Maugeri M, Mercalli L, Mestre O, Moisselin JM, Begert M, Müller-Westermeier G, Kveton V, Bochnicek O, Stastny P, Lapin M, Szalai S, Szentimrey T, Cegnar T, Dolinar M, Gajic-Capka M, Zaninovic K, Majstorovic Z, Nieplova E (2007) HISTALP—historical instrumental climatological surface time series of the Greater Alpine Region. International Journal of Climatology 27(1):17–46, https://doi.org/10.1002/joc.1377, URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.1377, _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/joc.1377, 27, 17, 46

Austria S (2017) Gemeindeverzeichnis herausgegeben von Statistik Austria. Tech. rep, Statistik Austria

Beguería S (2006) Changes in land cover and shallow landslide activity: a case study in the Spanish Pyrenees. Geomorphology 74(1–4):196–206. https://doi.org/10.1016/j.geomorph.2005.07.018

Bell R (2007) Lokale und regionale Gefahren- und Risikoanalyse gravitativer Massenbewegungen an der Schwäbischen Alb. PhD Thesis, Rheinischen Friedrich-Wilhelms-Universität Bonn, Bonn, URL http://hss.ulb.uni-bonn.de/2007/1107/1107.htm

Bell R, Petschko H, Röhrs M, Dix A (2012) Assessment of landslide age, landslide persistence and human impact using airborne laser scanning digital terrain models. Geografiska Annaler, Series A: Physical Geography 94(1):135–156

Beven KJ, Kirkby MJ (1979) A physically based, variable contributing area model of basin hydrology. Hydrological sciences 24:43–69

BMLFUW (2007) Hydrologischer Atlas Österreichs. Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft, Wien, In

Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MHH, White JSS (2009) Generalized linear mixed models: a practical guide for ecology and evolution. Trends Ecol Evol 24(3):127–135. https://doi.org/10.1016/j.tree.2008.10.008

Brabb EE (1984) Innovative approaches to landslide hazard and risk mapping. In: The 4th International Symposium on Landslides, The 4th International Symposium on Landslides, 1105, vol 1, pp 307–324

Brardinoni F, Slaymaker O, Hassan MA (2003) Landslide inventory in a rugged forested watershed: a comparison between air-photo and field survey data. Geomorphology 54(3–4):179–196. https://doi.org/10.1016/s0169-555x(02)00355-0

Brenning A (2005) Spatial prediction models for landslide hazards: review, comparison and evaluation. Nat Hazards Earth Syst Sci 5:853–862

Brenning A (2009) Benchmarking classifiers to optimally integrate terrain analysis and multispectral remote sensing in automatic rock glacier detection. Remote Sens Environ 113(1):239–247. https://doi.org/10.1016/j.rse.2008.09.005

Brenning A (2012) Spatial cross-validation and bootstrap for the assessment of prediction rules in remote sensing: the R package sperrorest. International Geoscience and Remote Sensing Symposium (IGARSS), pp 5372–5375, DOI https://doi.org/10.1109/IGARSS.2012.6352393

Brenning A, Trombotto D (2006) Logistic regression modeling of rock glacier and glacier distribution: topographic and climatic controls in the semi-arid Andes. Geomorphology 81(1):141–154. https://doi.org/10.1016/j.geomorph.2006.04.003

Budimir MEA, Atkinson PM, Lewis HG (2015) A systematic review of landslide probability mapping using logistic regression. Landslides 12(3):419–436. https://doi.org/10.1007/s10346-014-0550-1

Camilo DC, Lombardo L, Mai PM, Dou J, Huser R (2017) Handling high predictor dimensionality in slope-unit-based landslide susceptibility models through LASSO-penalized generalized linear model. Environ Model Softw 97:145–156. https://doi.org/10.1016/j.envsoft.2017.08.003

Cascini L (2008) Applicability of landslide susceptibility and hazard zoning at different scales. Eng Geol 102(3–4):164–177. https://doi.org/10.1016/j.enggeo.2008.03.016

Catani F, Lagomarsino D, Segoni S, Tofani V (2013) Landslide susceptibility estimation by random forests technique: sensitivity and scaling issues. Natural Hazards and Earth System Sciences 13(11):2815–2831, DOI https://doi.org/10.5194/nhess-13-2815-2013, URL https://www.nat-hazards-earth-syst-sci.net/13/2815/2013/

Chung CJF, Fabbri A (2003) Validation of spatial prediction models for landslide hazard mapping. Nat Hazards 30(3):451–472. https://doi.org/10.1023/B:NHAZ.0000007172.62651.2b

Coe JA, Michael JA, Crovelli RA, Savage WZ, Laprade WT, Nashem WD (2004) Probabilistic assessment of precipitation-triggered landslides using historical records of landslide occurrence, Seattle, Washington. Environ Eng Geosci 10(2):103–122. https://doi.org/10.2113/8.4.279

Conoscenti C, Rotigliano E, Cama M, Caraballo-Arias NA, Lombardo L, Agnesi V (2016) Exploring the effect of absence selection on landslide susceptibility models: a case study in Sicily, Italy. Geomorphology 261(Supplement C):222–235, DOI 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. Geoscientific Model Development Discussions 8:2271–2312. https://doi.org/10.5194/gmdd-8-2271-2015

Core Team R (2020) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, URL https://www.R-project.org/

Corominas J, Westen VC, Frattini P, Cascini L, Malet JP, Fotopoulou S, Catani F, MVD E, Mavrouli O, Agliardi F, Pitilakis K, Winter MG, Pastor M, Ferlisi S, Tofani V, Hervás J, Smith JT (2014) Recommendations for the quantitative analysis of landslide risk. Bull Eng Geol Environ 73(2):209–263. https://doi.org/10.1007/s10064-013-0538-8

Cruden DM, Varnes D (1996) Landslide types and processes. Landslide types and processes. In: turner AK, Schuster RL (eds) Landslides, investigation and mitigation. Transportation Research Board Special Report 247, Washington D.C, pp 36–75. ResearchGate 247:76

Dai FC, Lee CF (2002) Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. Geomorphology 42(3–4):213–228

Dikau R, Brunsden D, Schrott L, Ibsen M (eds) (1996) Landslide recognition. Identification, movement and causes. John Wiley & Sons Ltd, Chichester

Dikau R, Glade T (2003) Nationale Gefahrenhinweiskarte gravitativer Massenbewegungen. In: Liedtke H, Mäusbacher R, Schmidt KH (eds) Relief, Boden und Wasser, Nationalatlas Bundesrepublik Deutschland, vol 2. Spektrum Akademischer Verlag, Heidelberg, pp 98–99

Domínguez-Cuesta M, Bobrowsky PT (2017) Proposed landslide susceptibility map of Canada based on GIS. 4th World Landslide Forum, URL https://www.researchgate.net/publication/275963842_Proposed_Landslide_Susceptibility_Map_of_Canada_Based_on_GIS

Erener A, Düzgün HSB (2012) Landslide susceptibility assessment: what are the effects of mapping unit and mapping method? Environ Earth Sci 66(3):859–877. https://doi.org/10.1007/s12665-011-1297-0

European Environment Agency, EEA (2012) Corine Land Cover (CLC) 2012, Version 2020_20u1. URL https://land.copernicus.eu/pan-european/corine-land-cover/clc-2012?tab=metadata

Felicísimo NM, Cuartero A, Remondo J, Quirós E (2013) Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: a comparative study. Landslides 10(2):175–189. https://doi.org/10.1007/s10346-012-0320-1

Fell R, Corominas J, Bonnard C, Cascini L, Leroi E, Savage WZ (2008) Guidelines for landslide susceptibility, hazard and risk zoning for land-use planning. Eng Geol 102(3–4):99–111. https://doi.org/10.1016/j.enggeo.2008.03.014

Ferentinou M, Chalkias C (2013) Mapping mass movement susceptibility across Greece with GIS, ANN and statistical methods. In: Margottini C, Canuti P, Sassa K (eds) Landslide science and practice, Springer Berlin Heidelberg, pp 321–327, URL https://doi.org/10.1007/978-3-642-31325-7_42

Frattini P, Crosta G, Carrara A (2010) Techniques for evaluating the performance of landslide susceptibility models. Engineering Geology, URL 111(1):62–72. https://doi.org/10.1016/j.enggeo.2009.12.004 http://www.sciencedirect.com/science/article/pii/S0013795209003007

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(3):569–588. https://doi.org/10.5194/nhess-14-569-2014

Gaprindashvili G, Westen CJV (2015) Generation of a national landslide hazard and risk map for the country of Georgia. Nat Hazards 80(1):69–101. https://doi.org/10.1007/s11069-015-1958-5

Glade T, Crozier MJ (2005) A review of scale dependency in landslide hazard and risk analysis. In: Glade T, Anderson MG, Crozier MJ (eds) Landslide hazard and risk. Wiley, Chichester, pp 75–138

Glade T, Mergili M, Satler K (2020) ExtremA 2019. Aktueller Wissensstand zu Extremereignissen alpiner Naturgefahren in Österreich. Vienna University Press, 776 S

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

Graff JVD, Romesburg HC, Ahmad R, McCalpin JP (2012) Producing landslide-susceptibility maps for regional planning in data-scarce regions. Nat Hazards 64(1):729–749. https://doi.org/10.1007/s11069-012-0267-5

Günther A, Reichenbach P, Malet JP, Eeckhaut MVD, Hervás J, Dashwood C, Guzzetti F (2013) Tier-based approaches for landslide susceptibility assessment in Europe. Landslides 10(5):529–546. https://doi.org/10.1007/s10346-012-0349-1

Günther A, Van Den Eeckhaut M, Malet JP, Reichenbach P, Hervás J (2014) Climate-physiographically differentiated Pan-European landslide susceptibility assessment using spatial multi-criteria evaluation and transnational landslide information. Geomorphology, URL 224:69–85. https://doi.org/10.1016/j.geomorph.2014.07.011 http://www.sciencedirect.com/science/article/pii/S0169555X14003675

Guzzetti F (2005) Landslide hazard and risk assessment. PhD Thesis, Rheinische Friedrich-Wilhelms-Universität, Bonn

Guzzetti F, Carrara A, Cardinali M, Reichenbach P (1999) Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. Geomorphology 31(1–4):181–216

Guzzetti F, Mondini AC, Cardinali M, Fiorucci F, Santangelo M, Chang KT (2012) Landslide inventory maps: new tools for an old problem. Earth Sci Rev 112(1–2):42–66. https://doi.org/10.1016/j.earscirev.2012.02.001

Guzzetti F, Reichenbach P (1994) Towards a definition of topographic divisions for Italy. Geomorphology 11:57–74

Guzzetti F, Reichenbach P, Ardizzone F, Cardinali M, Galli M (2006) Estimating the quality of landslide susceptibility models. Geomorphology 81(1–2):166–184. https://doi.org/10.1016/j.geomorph.2006.04.007

Harp EL, Keefer DK, Sato HP, Yagi H (2011) Landslide inventories: the essential part of seismic landslide hazard analyses. Eng Geol 122(1):9–21. https://doi.org/10.1016/j.enggeo.2010.06.013

Hiebl J, Frei C (2016) Daily temperature grids for Austria since 1961—concept, creation and applicability. Theoretical and Applied Climatology 124(1):161–178. https://doi.org/10.1007/s00704-015-1411-4

Hong Y, Adler R, Huffman G (2007) Use of satellite remote sensing data in the mapping of global landslide susceptibility. Natural Hazards 43(2):245–256. https://doi.org/10.1007/s11069-006-9104-z

Hornich R, Adelwöhrer R (2010) Landslides in Styria in 2009. Geomechanics and Tunnelling 3(5):455–461

Hosmer DW, Lemeshow S (2000) Applied logistic regression, 2nd edn. Wiley series in probability and mathematical statistics - applied probability and statistics section. A Wiley-Interscience publication, Wiley, New York

Hungr O, Leroueil S, Picarelli L (2014) The Varnes classification of landslide types, an update. Landslides 11(2):167–194. https://doi.org/10.1007/s10346-013-0436-y

Hussin HY, Zumpano V, Reichenbach P, Sterlacchini S, Micu M, van Westen C, Balteanu 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

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

Kanungo D, Arora M, Sarkar S, Gupta R (2009) Landslide susceptibility zonation (LSZ) mapping-a review. South Asia Disaster Stud 2:81–105

Komac M (2006) A landslide susceptibility model using the analytical hierarchy process method and multivariate statistics in perialpine Slovenia. Geomorphology 74:17–28

Komac M, Ribicic M (2006) Landslide susceptibility map of Slovenia at scale 1 : 250,000. Geologija 49(2):295–309. https://doi.org/10.5474/geologija.2006.022

Krenmayr HG, Hofmann T, Mandl GW, Peresson H, Pestal G, Pistotnik J, Reitner J, Scharber S (2000) Rocky Austria. An illustrated earth history of Austria. Geological Survey of Austria, Vienna

Lee S, Choi J, Woo I (2004) The effect of spatial resolution on the accuracy of landslide susceptibility mapping: a case study in Boun, Korea. Geosci J 8(1):51–60. https://doi.org/10.1007/BF02910278

Lin L, Lin Q, Wang Y (2017) Landslide susceptibility mapping on a global scale using the method of logistic regression. Nat Hazards Earth Syst Sci 17(8):1411–1424. https://doi.org/10.5194/nhess-17-1411-2017

Liu C, Li W, Wu H, Lu P, Sang K, Sun W, Chen W, Hong Y, Li R (2013) Susceptibility evaluation and mapping of China's landslides based on multi-source data. Nat Hazards 69(3):1477–1495. https://doi.org/10.1007/s11069-013-0759-y

Lombardo L, Opitz T, Huser R (2018) Point process-based modeling of multiple debris flow landslides using INLA: an application to the 2009 Messina disaster. Stoch Env Res Risk A 32(7):2179–2198. https://doi.org/10.1007/s00477-018-1518-0

Malamud BD, Reichenbach P, Rossi M, Mihir M (2014) Report on standards for landslide modelling and terrain zonations. Tech. rep., KCL; King's College London, URL http://www.lampre-project.eu, [Online; accessed 2017-07-13]

Malamud BD, Turcotte DL, Guzzetti F, Reichenbach P (2004) Landslide inventories and their statistical properties. Earth Surface Processes and Landforms 29(6):687–711, URL ://000222253600003

Malet JP, Thiery Y, Hervás J, Günther A, Puissant A, Grandjean G (2008) Landslide susceptibility mapping at 1:1 M scale over France: exploratory results with a heuristic model. First French Conference on Landslides

Moosavi V, Niazi Y (2016) Development of hybrid wavelet packet-statistical models (WP-SM) for landslide susceptibility mapping. Landslides 13(1):97–114. https://doi.org/10.1007/s10346-014-0547-0

Murillo-García FG, Steger S, Alcántara-Ayala I (2019) Landslide susceptibility: a statistically-based assessment on a depositional pyroclastic ramp. J Mt Sci 16(3):561–580. https://doi.org/10.1007/s11629-018-5225-6

Nadim F, Kjekstad O, Peduzzi P, Herold C, Jaedicke C (2006) Global landslide and avalanche hotspots. Landslides 3(2):159–173. https://doi.org/10.1007/s10346-006-0036-1

Nefeslioglu HA, Duman TY, Durmaz S (2008) Landslide susceptibility mapping for a part of tectonic Kelkit Valley (eastern Black Sea region of Turkey). Geomorphology 94(3):401–418. https://doi.org/10.1016/j.geomorph.2006.10.036

Osborne JW (2015) Best practices in logistic regression. 55 City Road, London, URL http://methods.sagepub.com/book/best-practices-in-logistic-regression

Petley D (2012) Global patterns of loss of life from landslides. Geology 40(10):927–930. https://doi.org/10.1130/G33217.1

Petschko H, Bell R, Glade T (2016) Effectiveness of visually analyzing LiDAR DTM derivatives for earth and debris slide inventory mapping for statistical susceptibility modeling. Landslides 13(5):857–872. https://doi.org/10.1007/s10346-015-0622-1

Petschko H, Bell R, Leopold P, Heiss G, Glade T (2013) Landslide inventories for reliable susceptibility maps in lower Austria. In: Margottini C, Canuti P, Sassa K (eds) Landslide science and practice: volume 1: landslide inventory and susceptibility and hazard zoning, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 281–286

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(1):95–118. https://doi.org/10.5194/nhess-14-95-2014

Pourghasemi HR, Moradi HR, Fatemi Aghda SM (2013) Landslide susceptibility mapping by binary logistic regression, analytical hierarchy process, and statistical index models and assessment of their performances. Nat Hazards 69(1):749–779. https://doi.org/10.1007/s11069-013-0728-5

Pourghasemi HR, Rossi M (2016) Landslide susceptibility modeling in a landslide prone area in Mazandaran Province, north of Iran: a comparison between GLM, GAM, MARS, and M-AHP methods. Theoretical and Applied Climatology pp 1–25, DOI https://doi.org/10.1007/s00704-016-1919-2, 130, 609, 633

Regmi NR, Giardino JR, McDonald EV, Vitek JD (2014) A comparison of logistic regression-based models of susceptibility to landslides in western Colorado, USA. Landslides 11(2):247–262. https://doi.org/10.1007/s10346-012-0380-2

Reichenbach P, Busca C, Mondini AC, Rossi M (2014) The influence of land use change on landslide susceptibility zonation: the Briga catchment test site (Messina, Italy). Environ Manag 54(6):1372–1384. https://doi.org/10.1007/s00267-014-0357-0

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

Remondo J, González A, De Terán JRD, Cendrero A, Fabbri A, Chung CJF (2003) Validation of landslide susceptibility maps; examples and applications from a case study in northern Spain. Natural Hazards 30(3):437–449, URL http://springerlink.metapress.com/openurl.asp?genre=article&id=doi:10.1023/B:NHAZ.0000007201.80743.fc, 30, 437, 449

Sabatakakis N, Koukis G, Vassiliades E, Lainas S (2013) Landslide susceptibility zonation in Greece. Nat Hazards 65(1):523–543. https://doi.org/10.1007/s11069-012-0381-4

Schlögel R, Marchesini I, Alvioli M, Reichenbach P, Rossi M, Malet JP (2018) Optimizing landslide susceptibility zonation: effects of DEM spatial resolution and slope unit delineation on logistic regression models. Geomorphology 301:10–20. https://doi.org/10.1016/j.geomorph.2017.10.018

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

Schweigl J, Hervás J (2009) Landslide mapping in Austria. Tech. rep., European Commission Joint Research Centre Institute for Environment and Sustainability

Schwenk H (1992) Massenbewegungen in Niederösterreich 1953–1990. Jahrb Geol Bundesanst 135(2):597–660

Steger S, Brenning A, Bell R, Glade T (2016b) The propagation of inventory-based positional errors into statistical landslide susceptibility models. Nat Hazards Earth Syst Sci 16(12):2729–2745. https://doi.org/10.5194/nhess-16-2729-2016

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, Brenning A, Bell R, Petschko H, Glade T (2016a) 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, Glade T (2017) The challenge of "trivial areas" in statistical landslide suscep-tibility modelling. In: Advancing culture of living with landslides, Springer International Publishing, vol 2 Advances in Landslide Science (Proceedings of the 4th World Landslide Forum, May 29-June 2, Ljubljana), backup Publisher: Proceedings of the 4th World Landslide Forum

Steger S, Kofler C (2019) Statistical modeling of landslides: landslide susceptibility and beyond. In: Pourghasemi HR, Gokceoglu C (eds) Spatial modeling in GIS and R for earth and environmental sciences, Elsevier, pp. 519–546, dOI: https://doi.org/10.1016/B978-0-12-815226-3.00024-7

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. Science of The Total Environment, URL 776:145935. https://doi.org/10.1016/j.scitotenv.2021.145935 https://www.sciencedirect.com/science/article/pii/S0048969721010020

Steger S, Schmaltz E, Glade T (2020) The (f)utility to account for pre-failure topography in data-driven landslide susceptibility modelling. Geomorphology, URL 354:107041.

https://doi.org/10.1016/j.geomorph.2020.107041 https://www.sciencedirect.com/science/article/pii/S0169555X20300118

Sterlacchini S, Ballabio C, Blahut J, Masetti M, Sorichetta A (2011) Spatial agreement of predicted patterns in landslide susceptibility maps. Geomorphology 125(1):51–61. https://doi.org/10.1016/j.geomorph.2010.09.004

Süzen ML, Kaya B (2012) Evaluation of environmental parameters in logistic regression models for landslide susceptibility mapping. International Journal of Digital Earth 5(4):338–355

Terhorst B, Damm B (2009) Slope stability and slope formation in the Flysch zone of the Vienna Forest (Austria). Journal of Geological Research pp:1–10

Tien Bui D, Pradhan B, Lofman O, Revhaug I (2012) Landslide susceptibility assessment in Vietnam using support vector machines, decision tree, and Naïve Bayes models. URL https://www.hindawi.com/journals/mpe/2012/974638/, publication Title: Mathematical Problems in Engineering

Tilch N (2009) Gravitative Massenbewegungen in der Katastrophenregion Klingfurth (Walpersbach, Südliches Niederösterreich) im Juni 2009 - Erkundungsergebnisse und eine erste Abschätzung des rutschungsinduzierten Gefahrenpotentials. In: Geoforum Umhausen, 11. Geoforum Umhausen 15.-16.10.09, Niederthai, p http://www.geologie.ac.at/pdf/Poster/poster_2009_geoforum_tilch.pdf

Trigila A, Frattini P, Casagli N, Catani F, Crosta G, Esposito C, Iadanza C, Lagomarsino D, Mugnozza GS, Segoni S, Spizzichino D, Tofani V, Lari S (2013) Landslide susceptibility mapping at national scale: the Italian case study. In: Margottini C, Canuti P, Sassa K (eds) Landslide science and practice, Springer Berlin Heidelberg, pp 287–295, URL https://doi.org/10.1007/978-3-642-31325-7_38

van Den Eeckhaut M, Hervas J, Jaedicke C, Malet JP, Montanarella L, Nadim F (2012) Statistical modelling of Europe-wide landslide susceptibility using limited landslide inventory data. Landslides published online:1–13, DOI https://doi.org/10.1007/s10346-011-0299-z, 9, 357, 369

van Den Eeckhaut M, Reichenbach P, Guzzetti F, Rossi M, Poesen J (2009) Combined landslide inventory and susceptibility assessment based on different mapping units: an example from the Flemish Ardennes, Belgium. Nat Hazards Earth Syst Sci 9(2):507–521. https://doi.org/10.5194/nhess-9-507-2009

van Westen CJ, Castellanos E, Kuriakose SL (2008) Spatial data for landslide susceptibility, hazard, and vulnerability assessment: an overview. Engineering Geology, URL 102(3):112–131. https://doi.org/10.1016/j.enggeo.2008.03.010 https://www.sciencedirect.com/science/article/pii/S0013795208001786

Weber L (1997) Geologische Karte 1:500.000 und die Ebene "Postobereozäne Becken und Quartär 1:500.000"

Weiss A (2001) Topographic position and landforms analysis. ESRI User Conference, San Diego, CA

Wilde M, Günther A, Reichenbach P, Malet JP, Hervás J (2018) Pan-European landslide susceptibility mapping: ELSUS version 2. Journal of Maps 14(2), DOI https://doi.org/10.1080/17445647.2018.1432511, https://doi.org/10.1080/17445647.2018.1432511, publisher: Taylor & Francis

Wu X, Chen X, Zhan FB, Hong S (2015) Global research trends in landslides during 1991–2014: a bibliometric analysis. Landslides 12(6):1215–1226. https://doi.org/10.1007/s10346-015-0624-z

Zêzere JL, Pereira S, Melo R, Oliveira SC, Garcia RAC (2017) Mapping landslide susceptibility using data-driven methods. Sci Total Environ 589:250–267. https://doi.org/10.1016/j.scitotenv.2017.02.188

Zuur AF, Leno EN, Walker NJ, Saveliev AA, Smith GM (2009) Mixed effects models and extensions in ecology with R. Springer, New York, NY

**P. Lima** (✉) · **T. Glade**
Department of Geography and Regional Research,
University of Vienna,
Vienna, Austria
Email: pedro.lima@univie.ac.at

**S. Steger**
Eurac Research, Institute for Earth Observation,
Bozen, Italy