Nat. Hazards Earth Syst. Sci. Discuss., 1, 1001–1050, 2013
www.nat-hazards-earth-syst-sci-discuss.net/1/1001/2013/
doi:10.5194/nhessd-1-1001-2013

Natural Hazards and Earth System Sciences
Open Access
Discussions

Assessing the quality of landslide susceptibility maps – case study Lower Austria

H. Petschko[1], A. Brenning[2], R. Bell[1], J. Goetz[1,2], and T. Glade[1]

[1]Department of Geography and Regional Research, University of Vienna, Vienna, Austria
[2]Department of Geography and Environmental Management, University of Waterloo, Ontario N2L 3G1, Canada

Received: 27 February 2013 – Accepted: 12 March 2013 – Published: 10 April 2013

Correspondence to: H. Petschko (helene.petschko@univie.ac.at)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Landslide susceptibility maps are helpful tools to identify areas which might be prone to future landslide occurrence. As more and more national and provincial authorities demand for these maps to be computed and implemented in spatial planning strategies, the quality of the landslide susceptibility map and of the model applied to compute them is of high interest. In this study we focus on the analysis of the model performance by a repeated k-fold cross-validation with spatial and random subsampling. Furthermore, the focus is on the analysis of the implications of uncertainties expressed by confidence intervals of model predictions. The cross-validation performance assessments reflects the variability of performance estimates compared to single hold-out validation approaches that produce only a single estimate. The analysis of the confidence intervals shows that in 85 % of the study area, the 95 % confidence limits fall within the same susceptibility class. However, there are cases where confidence intervals overlap with all classes from the lowest to the highest class of susceptibility to landsliding. Locations whose confidence intervals intersect with more than one susceptibility class are of high interest because this uncertainty may affect spatial planning processes that are based on the susceptibility level.

1 Introduction

Landslides occur in mountainous as well as in hilly or coastal regions worldwide and have often been an underestimated hazard. In general, people and governing authorities are not sufficiently aware of the potential locations and consequences of landslides (Hervás, 2003). However, in Austria authorities and residents have become more aware of landslides hazards because of recent major landslide events, which affected many residents, caused significant damage to infrastructure and private properties and were reported and discussed on the local media (Damm et al., 2010). These include events in August 2005 in Gasen and Haslau and incidents in 2009 in the district of Feldbach and in the province of Lower Austria where about 4000 landslides occurred in total (Schwarz and Tilch, 2008; Hornich and Adelwöhrer, 2010; Abteilung Feuerwehr und Zivilschutz, Amt der NÖ Landesregierung, 2010; BMLFUW, 2010). While in the past in Austria practices of post-disaster recovery and reconstruction were applied, recently more and more national and provincial authorities demand for pre-disaster mitigation tools which help to prevent future damage caused by natural hazards such as floods and landslides. Aiming at avoidance of the hazard, zonation plans can be facilitated in prospective land use planning to prevent development in undesirable locations or undesirable types of development (Schwab et al., 2005). In this context landslide susceptibility maps have proven to be a powerful tool, as they give coherent information on the spatial probability on where landslides, or landslide scarps, might occur (Brabb, 1984; Glade et al., 2005; Guzzetti et al., 2000, 2006; Varnes, 1984). The term landslide susceptibility is hereby defined by Brabb (1984) as the likelihood of a landslide occurring in an area with given local terrain attributes.

The central assumption at modelling of landslide susceptibility is the law of uniformity "the past and the present are keys to the future" which is based on the concept of uniformitarianism of James Hutton (1726–1797) (Orme, 2002). This is applied to statistically model the possible future location of landslides using information on the location and local terrain attributes of past landslides usually documented in a landslide inventory (Carrara et al., 1995). Statistical landslide susceptibility models are particularly useful for modelling large areas to get an overview of which slopes or slope sections might be prone to landslides in future. The number of different statistical modelling methods applied in the context of landslide susceptibility has risen manifold recently. Detailed reviews and comparison of different models can be found, among others, in Guzzetti et al. (1999); Dai et al. (2002), Brenning (2005), Glade and Crozier (2005), Guzzetti (2005), Rossi et al. (2010) and Vorpahl et al. (2012). Summarizing these methods it can be stated that models using machine learning algorithms tend to have overfitting, while linear models might not be flexible enough to portray possible nonlinearity in the relationship between the occurrence of landslides and the explanatory variables (Brenning, 2005; Goetz et al., 2011).

In earlier studies single hold-out model performance measures were derived using one training and independent test set (i.e. Chung and Fabbri, 1999, 2003; Fabbri et al., 2003; Remondo et al., 2003; Brenning, 2005; Beguería, 2006; Frattini et al., 2010; Rossi et al., 2010). However, we identified that there is a need for a more transparent and reliable estimation of the model performance and spatial transferability and propose the assessment by a repeated *k*-fold cross-validation with *k* training and test sets (Brenning, 2012a, b). This assessment results in a range of possible AUROC values instead of only one single "random" AUROC value obtained using one test set. It is therefore more reliable than traditional methods. Regarding the influence of the model uncertainty on the resulting map, previous research (e.g. Guzzetti et al., 2006; Van den Eeckhaut et al., 2009; Rossi et al., 2010; Sterlacchini et al., 2011) did not give an estimate on the spatial implications of uncertainties in the prediction on the final classified map and ways of communicating them.

Based on these findings our research objectives are to assess the spatial and non-spatial transferability of a generalized additive model (GAM) of landslide susceptibility in Lower Austria with a repeated *k*-fold cross-validation and to analyse the sensitivity of the classified map to the uncertainties in the predicted susceptibility probabilities.

2 Considerations on the quality of landslide susceptibility maps

As it comes to the application of the landslide susceptibility maps, which is associated with constructing a decisive reality (Egner and Pott, 2010) for the municipality and the land use planners, a detailed and transparent assessment of its quality is necessary. Since the application of the landslide susceptibility model in this study is actually planned for implementation by municipal authorities of Lower Austria, our main focus is to identify different aspects of quality and to analyse them by statistical methods. A short review of different aspects of quality in landslide susceptibility assessments is provided as these are a basis for identifying the research gap and objective of this study.

Arising from missing knowledge and imperfect understanding of the complexity of hazardous processes such as landslides and their natural variability, landslide susceptibility/hazard assessments are always prone to uncertainties (Ardizzone et al., 2002; Kunz et al., 2011). Besides these “a priori” uncertainties which cannot be neglected but also not quantified there are more aspects that influence the quality of a landslide susceptibility map. This very broad term of quality can be interpreted in several ways and on several stages of the process of preparing a landslide susceptibility map as described, amongst others, by Carrara et al. (1995), Ardizzone et al. (2002), and Guzzetti et al. (2006). A summary of these (refer to the following points 1–3) and some of our own additions (points 3 and 4) are presented below. It should be understood that all aspects together are resulting in the quality of the final susceptibility map:

(1) *Quality of the input data:* achieving a good quality of the final landslide susceptibility map starts with the quality of the input data set for the modelling. Besides the geomorphological relevance, the spatial resolution and accuracy of the geo-environmental as well as the landslide inventory data is also important. Furthermore, possible incompleteness, not only in a sense of full spatial coverage but also of general availability of important thematic information on (predisposing or preparatory) factors determining the landslide susceptibility influences the quality (Carrara et al., 1999; Ardizzone et al., 2002). Estimation on the completeness of the landslide inventory and details on the collecting and mapping method giving information on the accuracy and the location of the landslide point/line/polygon (main scarp or entire landslide body) are very important. Both influence the further usage of the input data set and the feasible interpretation of the maps substantially although it is difficult to explicitly incorporate these influences in the subsequent hazard and risk assessment (Ardizzone et al., 2002).

(2) *Quality of the statistical model:* in the modelling stage the quality can be determined by the assessment of the model performance, transferability and its geomorphologic plausibility. The prediction skills (Guzzetti et al., 2006) can be analysed

with several quantitative model performance measures and estimation methods but also with qualitative methods. Common performance measures are success/prediction rate (Chung and Fabbri, 2003, 2008), confusion matrix or error rates (Beguería, 2006; Brenning, 2005), or cost curves (Frattini et al., 2010). Among the performance estimation methods hold-out validation or cross-validation with determining the area under receiver operating characteristic curve (AUROC) value and ROC plots (Beguería, 2006; Brenning, 2005) is usually applied. Statistical estimation methods further provide a means for assessing the non-spatial transferability of a model onto a different, independent random sample, and the spatial transferability into a spatially separate area. Spatial transferability refers to the capability of the model to generalize empirical relationships learned on a training data set, and to transfer these relationships to (usually adjacent) regions without major loss in predictive performance (Brenning, 2005; Von Ruette et al., 2011). Qualitative methods analyse the geomorphologic plausibility of the map (Demoulin und Chung, 2007; Bell, 2007).

(3) *Quality in terms of transparent analysis of uncertainties:* quality can also be defined by the analysis of uncertainties involved in the final susceptibility map. Besides the uncertainties resulting from the input data (refer to point 1) also the modelling itself introduces uncertainties. The result of statistical modelling methods, such as logistic regression or generalized additive models, is an estimated conditional mean value of the predicted probability (Hosmer and Lemeshow, 2000) which is used for showing the susceptibility of each cell (slope/terrain unit) of a map. Therefore, there is an uncertainty or possible range (as determined by the standard error of the predicted probabilities) in the estimates of the probability of each cell of the entire map (Guzzetti et al., 2006). When landslide susceptibility maps are used for planning purposes the analysis and presentation of this standard error is relevant (Guzzetti et al., 2006; Rossi et al., 2010). Especially the way they affect the appearance of the classified landslide susceptibility map is of interest in the planning context as this might result in overlaps of different susceptibility classes.

NHESSD
1, 1001–1050, 2013

(4) *Quality achieved by communication and presentation of uncertainties:* finally we propose that clear and transparent end-user oriented documentation on all these aspects of quality contributes to a high quality of a landslide susceptibility map. This is of high importance for the acceptance and implementation of landslide susceptibility maps in land use planning (Guzzetti et al., 2006). There is a need to communicate the research results and their quality with appropriate explanations for the local officials, environmental managers and the public to raise awareness and knowledge on it which leads to an easier understanding and incorporation of the results into the decision-making process (Knuepfer and Petersen, 2002; Rogers, 2006; Brierley, 2009). In particular, the visualization of some aspects of the quality of landslide susceptibility maps such as the uncertainty of the probability value can enhance the communication among experts and decision-makers to facilitate informed decisions (Kunz et al., 2011).

3 Lower Austria – a heterogeneous province

Austria's north-eastern province Lower Austria covers a total area of 19 177 km^2, 15 850 km^2 of which constitute our study area (Fig. 1). Three districts (Gmünd, Mistelbach, Gänserndorf) in the north of the province are not included in this study as the topography and lithology show characteristics that have not been prone to slides in the past.

The heterogeneity of the study area results from the high number of different lithological units with a large variety of associated geotechnical and topographical characteristics. The lithology in the study area is geotechnically summarized according to predominant material types into 20 groups (Fig. 1, Table 1). The predominant material types range from gravel, sand, loess and loam in the alluvial deposits and fluvial terraces, to marl with high amount of silt in the Molasse Zone and Schlier Zone, to sandstone, marlstone of the Rheno-danubian Flysch Zone and of the Mélange Zone, to limestone, dolostone and the ingeous rocks of the Austroalpine Unit and to different gneisses and granites in the Bohemian Massif in the north(-west) of Lower Austria.

NHESSD
1, 1001–1050, 2013

In terms of topographic characteristics, the median slope angle ranges from a minimum of $< 1^{\circ}$ (alluvial deposits) to a maximum of 27° (Austroalpine Unit with dolostone) (Table 1).

Land use patterns also differ among lithological groups and topography. The predominant land uses in the province (40 % agricultural land, 40 % forest, 13 % grassland) are furthermore distributed unevenly across the province. While the relatively flat north-east of the province and areas along and south of the Danube are predominantly used as agricultural land, the steeper slopes in the south and south-west are mainly forest-covered (coniferous and deciduous forest) (Eder et al., 2011). The spatial distribution of the mean annual precipitation rate shows a gradient between the north(-east) (400–500 mm) and the south(-west) (1600–1700 mm) (Hydrographischer Dienst des Landes Niederösterreich, 2011).

In the province a full range of landslide types with rock falls, earth slides, debris slides and debris flows (classified according to Cruden and Varnes, 1996; Dikau et al., 1996) is present. While rock fall and debris flows can mainly be found in the south, in the Austroalpine Unit, earth and debris slides occur all over the province with different density (Table 1), size and depth. In this study we focus on earth and debris slides (recent examples in Fig. 2).

The slides in Lower Austria are mainly triggered by heavy rainfall or rapid snow melt events (Schwenk, 1992; Schweigl and Hervás, 2009). The landslide density, i.e. the number of landslides per square kilometre, shows a maximum of 5.5 km^{-2} in the Mélange Zone. Furthermore, the Rheno-danubian Flysch Zone and the Zone of Molasse with Schlier show a very high landslide density with more than four landslides per square kilometre. The lithological material and topographic characteristics of these units result in a high susceptibility to earth and debris slides (Gottschling, 2006; Wessely, 2006).

4 Data sources and preparation

Practical challenges in this study arise from the size of the study area and the intended output map scale of 1 : 25 000. The size of the study area brings along some limitations regarding the availability of data sources that offer a full spatial coverage and a reasonable map scale. However, an airborne laser scanning digital terrain model (ALS-DTM) with a resolution of 1 m × 1 m, acquired between 2006 and 2009, was available. ALS-DTMs are very useful for mapping landslides and representing the morphology of the study area, even in forest areas (Van den Eeckhaut et al., 2007a; Razak et al., 2011).

We decided on an output resolution of the landslide susceptibility maps of 10 m × 10 m as we wanted to take advantage of the high resolution of the topographic data given by the ALS-DTM. All data on the response variables were resampled to a 10 m × 10 m resolution for modelling purposes while being aware that this artificial improvement of the data resolution does not increase the data accuracy (details on the scale and resolution of the data is given in Sect. 4.1). By doing this we can use important details in the topographic information (from the ALS-DTM) for the susceptibility modelling while the soil properties and geological information are still included on a reasonable scale. In the following sections details on the type, source and data preparation of the response variable and the explanatory variables used for the modelling are given.

4.1 Response variable

The landslide inventory originates from previous research where earth and debris slides were mapped on the basis of a high resolution ALS-DTM (Petschko et al., 2010, 2013; Glade et al., 2012). This inventory consists of point locations representing each landslide's main scarp and includes 12 889 earth and debris slides. Considering the large study area a point inventory was preferred over a polygon inventory especially due to the increased mapping efficiency but also due to the avoidance of uncertainty in the delineation of landslide polygons (Van den Eeckhaut et al., 2006; Petschko et al., 2013). Comparisons between the use of point inventories with different point locations (e.g. one point for each main scarp or entire landslide) representing landslides showed very small differences in the resulting susceptibility maps and in the predictive model performances (Petschko et al., 2013). However, the full landslide with scarp, transportation and accumulation area had to be discernible in the ALS-DTM hillshade in order for it to be included in the inventory.

The mapping focussed on distinct and easily detectable morphologies that remain visible after the occurrence of a landslide (McCalpin, 1984). Very old landslides, or landslides that are not visible in the available imagery any more, are missing from the inventory (Bell et al., 2012). In areas where the human impact (depending on the land use type; e.g. farming, planation) is very high the landslides may not be visible anymore within a few years only (Bell et al., 2012). However, the full extent of the incompleteness of the inventory obtained in the study area remains unknown.

4.2 Explanatory variables

Geomorphic meaningful explanatory variables have been derived from the high resolution ALS-DTM, from raster data on soil properties (50 m × 50 m, Eder et al., 2011) and from vector data on tectonic lines in the study area (1 : 200 000, Kurz, 2012) as described in more detail in the following.

Furthermore, data on land cover were available for this study. However, the mapped landslides are of unknown age, and no information is available on historical land cover changes during the time period covered by the landslide inventory, which would be required to identify differences in susceptibility among different land uses. Because of this we decided not to use present-day land cover in the modelling as it might not represent the respective land cover at the time of the event (Petschko et al., 2013).

Several terrain attributes were derived from the ALS-DTM as proxies for geomorphic and hydrological processes. We used SAGA GIS (Conrad, 2007) to calculate slope angle (°), slope aspect (°, transformed using the sine and cosine representing the east versus west and the north versus south, Brenning, 2009), overall curvature (all calculated with second-degree polynomial approximation; Zevenbergen and Thornes, 1987), a topographic wetness index (SAGA wetness index of Boehner et al., 2002), catchment height, catchment area (all calculated with multiple flow direction algorithm, Freeman, 1991; Quinn et al., 1991) and convergence indices, calculated with 10 m and 50 m radius respectively to represent morphological changes on different scales. Data preparation was performed within the R package RSAGA (Brenning, 2011).

Information on the soil properties is also very important as they have an effect on the infiltration capacity and water storage in the soil, which ultimately influences the disposition to landslides (Crozier, 1986). From the soil dataset of Eder et al. (2011) we extracted four gridded variables representing saturated water conductivity (mm d^{-1}; average value of the top 20 cm, and minimum value within 100 cm profiles) and void space (%; average value of the top 20 cm, and average of 100 cm profiles) which may be considered as a proxy for the infiltration capacity.

Available tectonic vector data includes fault lines and nappe boundaries. The proximity to a fault line might refer to the occurrence of weakened material which has already been strongly tectonically influenced and reworked. This material shows lower shear strength and may therefore be more prone to landslides (Crozier, 1986). Furthermore, landslides might occur with higher density close to nappe boundaries as these indicate a distinct difference in the material and permeability. At the nappe boundary of the Austroalpine Unit with limestone overlaying the Rheno-danubian Flysch Zone, for example, many landslides occurred in the past. This might be related to the difference between mainly limestone with high water permeability on top of denser sandstones and marlstones of the Flysch Zone resulting in increased soil water availability in the boundary zone and presence of boundary springs (Schnabel, 1985). As both tectonic features may have different relationships with the disposition to landslides we decided to derive for each type a grid of the Euclidean distance to the lines as an additional explanatory variable.

NHESSD
1, 1001–1050, 2013

The lithological map (1 : 200 000, Schnabel, 2002, simplified by the Austrian Institute of Technology) is not used as explanatory variable, but for partitioning the study area into more homogenous modelling domains (refer to Sect. 5.1).

5 Methods

5.1 Modelling of heterogeneous areas

The study area's high heterogeneity in its geotechnical and topographical characteristics has to be taken into account in the landslide modelling to provide a susceptibility map which allows for the comparison of susceptibilities across the entire study area (Lee et al., 2008). In a previous study this heterogeneity and a new study design meeting the characteristics of the study area and its landslide susceptibility was analysed in a test area in Lower Austria (Petschko et al., 2012). This new study design includes dividing the study area in more homogeneous modelling domains according to the geotechnical and topographical conditions of the lithological units. Consequently, separate susceptibility models are fitted and assessed within each of these domains. This approach avoids the use of interaction terms to represent lithology-dependent processes and preparatory factors, and thus facilitates easier interpretation of the models. The use of different models for the distinct lithological units is also expected to achieve an improved overall predictive performance because the models are able to account for the distinct geotechnical and topographic characteristics of each lithological zone (Petschko et al., 2012).

Our study area is therefore divided into 16 homogenous modelling domains based on the lithological map. The analysed homogenous modelling domains are formed by merging the units with no observed landslides with geotechnically similar units (Table 1). Although the unit of Fault Zone within the Bohemian Massif is referred to as a tectonic fault zone, Kurz (2012) states that no major fragmentation can be expected in this area, consequently the geotechnical characteristics still resemble the surrounding Bohemian Massif lithological unit.

The final susceptibility map of the province is obtained by merging individual susceptibility maps. Therefore, the predicted landslide probability derived in the different modelling domains has to be adjusted according to the general tendency to landslides in the domain (refer to Sect. 5.2).

5.2 Generalized additive model

Generalized additive models (GAM, Hastie and Tibshirani, 1990) have recently been introduced in landslide susceptibility modelling (Brenning, 2008; Jia et al., 2008; Park and Chi, 2008; Goetz et al., 2011; Vorpahl et al., 2012). These models represent an extension to generalized linear models (GLMs), such as logistic regression, as a GAM can model nonlinear relationships between response and explanatory variables (Hastie and Tibshirani, 1990). Among the currently available methods for landslide susceptibility modelling a GAM shows a compromise between the flexibility of machine learning algorithms, the smooth representation which results from GLMs such as logistic regression and still gives the opportunity of a transparent and easy interpretable model (Brenning, 2008; Goetz et al., 2011). Furthermore, Goetz et al. (2011) showed that compared to a GLM the GAM is able to better reflect the nonlinear response of slope stability to changing site conditions.

The basic idea of a GAM is to replace the linear function of each covariate as used in a GLM with an empirically fitted smooth function to "let the data show the appropriate functional form" (Hastie and Tibshirani, 1990). Thus a GAM allows the combination of linear and nonlinear smoothing functions in an additive manner to describe the relationship between explanatory and response variables. In the case of the logistic additive model for binary (presence/absence) response variables, the response variable is not modelled directly, but using the logit of the occurrence probability $p(\mathbf{x})$ conditional on explanatory variables $\mathbf{x} = (x_1, \ldots, x_m)^T$ (Hastie and Tibshirani, 1990):

$$\text{logit}(\boldsymbol{x}) = \log(p(\boldsymbol{x})/(1 - p(\boldsymbol{x}))) = \beta_0 + \beta_1 f_1(x_1) + \ldots + \beta_m f_m(x_m) \quad (1)$$

where the functions f_i are nonparametric smoothers. The quantity

$$\text{odds}(\boldsymbol{x}) = p(\boldsymbol{x}) / (1 - p(\boldsymbol{x})) \quad (2)$$

is referred to as the odds. Thus, the logistic GAM is additive at the logit level, but increases in f_i have a multiplicative effect on the odds.

We use a combined backward and forward stepwise variable selection based on the Akaike Information Criterion (AIC) which measures the goodness-of-fit while penalizing for model complexity to obtain a parsimonious model that explains the occurrence of landslides almost as well as larger, more complex models (Akaike, 1974). Smaller models help to keep the estimated coefficient standard errors small and prevent the model from overfitting, which occurs especially when the number of variables in the model is large relative to the number of landslide points (Hosmer and Lemeshow, 2000). Overfitting means that an algorithm or model performs very well on the available training data to which it is fitted, but poorly on future or new test data and therefore produces unreliable predictions (Hand, 1997; Hosmer and Lemeshow, 2000).

Our study design is a case-control study with the mapped landslide points as cases and randomly selected non-landslide points as controls. The landslide susceptibility maps are derived for each modelling domain from a model (GAM_1–GAM_{16}) using all landslide points while the model performance is assessed using cross-validation, i.e. subsamples of the landslide points (Sect. 5.3). In both cases we use equal numbers of cases and controls (1 : 1), which means that within each homogeneous modelling domain, the landslide locations are matched to an equal number of randomly selected non-landslide locations. The sampling rates therefore vary among the domains depending on landslide density (Table 1).

To combine predictions of landslide probability produced by the models in the different modelling domains into one susceptibility map, it is necessary to adjust each

NHESSD
1, 1001–1050, 2013

Assessing the quality of landslide susceptibility maps of Lower Austria

H. Petschko et al.

model's raw predictions based on the corresponding sampling rate. The unadjusted predictions odds($\boldsymbol{x}$) of a model that is based on training data with a sampling rate

$$\tau_0 = \text{number of slide points}/\text{number of non-slide points} \tag{3}$$

for non-landslide points and a sampling rate $\tau_1 = 1$ for landslide points are transformed using

$$\text{odds}^*(\boldsymbol{x}) = \tau_0/\tau_1 \times \text{odds}(\boldsymbol{x}) \tag{4}$$

where odds$^*(\boldsymbol{x})$ gives the adjusted odds (Scott and Wild, 1986; Hosmer and Lemeshow, 2000). The adjusted probability $p^*(\boldsymbol{x})$ is calculated from odds$^*(\boldsymbol{x})$ which is therefore comparable among different modelling domains. GAM modelling is performed with the gam package within the open-source statistical software R (Hastie, 2011; R Development Core Team, 2011).

The predicted probabilities of landslide susceptibility were further classified into discrete classes of low, medium and high susceptibility based on thresholds related to the percentage of observed landslide points falling within each susceptibility class. This is a result of testing different thresholds and checking them in the field according to the best geomorphic and planning plausibility. A threshold for the low susceptibility class was selected so that 5 % of the observed landslides having the lowest predicted probabilities would fall within this class. Additional 25 % of the observed landslides would fall within the medium susceptibility class, while the high susceptibility class contained the 70 % of the landslide points with the highest predicted probabilities. The corresponding class thresholds of 0.00209 and 0.0141 were also applied to classify the maps corresponding to the approximate upper and lower confidence limits calculated for each grid cell at the 95 % confidence level (refer to Sect. 5.4).

5.3 Spatial and non-spatial cross-validation

The assembly of the test data for performance assessment can be achieved in three ways (1) random subsampling, (2) spatial subsampling, and (3) temporal partitioning of

NHESSD
1, 1001–1050, 2013

the landslide data (Chung and Fabbri, 2003). As there is no information on the landslide age or time of occurrence required for temporal partitioning, we focus on testing random and spatial subsampling of cases and controls within each homogeneous modelling domain. Consequently, the spatial transferability is not evaluated across modelling domains, but only within each modelling domain where geotechnical and topographic conditions are homogeneous and thus are expected to be comparable.

We use non-spatial and spatial k-fold cross-validation to assess each model's predictive performance as a measure of non-spatial and spatial transferability. In k-fold cross-validation, k (here: $k = 5$) randomly or spatially selected disjoint subsamples, or folds, are derived. The model is fitted k times on the combined data of $k-1$ folds and tested on the data of the remaining fold by applying the fitted model to the test fold and calculating the performance measure. We use random subsampling for traditional non-spatial cross-validation and spatial subsampling based on k-nearest-neighbour classification of point coordinates for spatial cross-validation (Ruß and Brenning, 2010), which assesses the model's non-spatial and spatial transferability.

In order to obtain results that are independent of a particular partitioning, cross-validation is repeated r times (here: $r = 20$), and the median and interquartile range of the 20 outcomes are calculated. This results in 100 different estimates of the performance measures.

We use the AUROC as a performance measure, which is derived by comparing the sensitivity of a model (proportion of true positives) to the specificity (more precisely, 1 – specificity, or false positives rate) (Hosmer and Lemeshow, 2000). The AUROC takes values between 0.5 and 1 and measures the model's ability to discriminate landslide and non-landslide points (Hosmer and Lemeshow, 2000). Values approaching 1 show perfect discrimination; however, this may also indicate overfitting (Brenning, 2005; Guzzetti et al., 2006).

5.4 Transferability and thematic consistency indices

We define a transferability index by adjusting the estimated interquartile range (IQR) for the contribution of test-set estimation to sampling variability. For this purpose, we calculate the approximate standard error SE of the AUROC estimator on a test set of n landslide and n non-landslide samples using Eq. (1) of Hanley and McNeil (1982). Since the IQR is approximately equal to 1.35 times the corresponding standard deviation under normal distribution, we use the following equation to remove the contribution of SE to the IQR of the AUROC, and refer to this quantity as the transferability index T:

$$T = \sqrt{\mathrm{IQR}^2 - (1.35 \cdot \mathrm{SE})^2} \tag{5}$$

when slightly negative terms occurred under the square root due to the approximation used, these values were replaced by a value of zero.

A higher value of T indicates a greater variation in predictive performance among models fitted to cross-validation training partitions, which can be interpreted as a poorer transferability. We refer to the transferability calculated by spatial (non-spatial) cross-validation as spatial (non-spatial) transferability T_{sp} (T_{nsp}).

Based on the 100 models fitted on the different cross-validation training sets within each modelling domain using stepwise variable selection we assess the importance of each variable in predicting landslide susceptibility using its selection frequency (Goetz et al., 2011). This variable-selection frequency can also be interpreted as a proxy for the thematic consistency of the model (Guzzetti et al., 2006). To formalize this concept, we introduce a thematic consistency index that measures the agreement between variable choices among cross-validation repetitions. In analogy to the Gini impurity index used in classification (Hand, 1997), we define the consistency index C by

$$C = (p_1(1 - p_1) + \ldots + p_m(1 - p_m)) / (0.25 \cdot m) \tag{6}$$

where p_i is the proportion of models that include the i-th predictor variable out of m predictors. The consistency index is calculated for each modelling domain and for spatial

(C_{sp}) and non-spatial cross-validation (C_{nsp}). Good thematic consistency is achieved when each variable has a selection frequency either close to 0 % or 100 %, resulting in a $p_m(1-p_m)$ value near 0. Therefore, a low value of C indicates a strong consistency among models. Selection frequencies around 50 % indicate weak thematic consistency and produce $p_m(1-p_m)$ values of up to 0.25 m and a maximum C value of 1. This and the associated AUROC variation reflect the model's sensitivity of the model to sampling variation.

Spatial and non-spatial cross-validation is also applied to test the effects of reducing the sample size on the interquartile range of the AUROC values and the thematic consistency within one domain. This is performed within the modelling domain Flysch Zone. The cross-validation is applied 9 times (with 5 folds and 20 repetitions each) while the sample size of the training sample is reduced from 12562 to 6400, 3200, 1600, 800, 400, 200, 100 and to 50 landslide and non-landslide samples. The test sample is determined constant with 2000 landslide and non-landslide sample points. Spatial and non-spatial cross-validation are performed with the sperrorest package in R (Brenning, 2012b).

5.5 Spatially varying prediction uncertainties

The basis for the visualisation of the landslide susceptibility map is the predicted probability of the occurrence of landslides of each grid cell which is computed from the predicted logit. These predictions are subject to uncertainty due to sampling variability and model error, which can be expressed by the standard error of model predictions. This standard error further provides a means to determine approximate upper and lower confidence limits for the predicted logit and ultimately the predicted probability. These limits define an interval within which the true logit or probability of sites with given values of the explanatory variables is located with the chosen confidence of, for example, 95 % (Hosmer and Lemeshow, 2000). In other words we have strong confidence that the true probability of landslide occurrence at a given type of location is within the confidence interval, but we would hesitate to claim that the true probability falls within any narrower range of values within the confidence interval.

In this study confidence interval estimates for the predicted logit and probability are of special interest in order to assess the implications of spatially varying uncertainties for the interpretation of the final classified susceptibility map. Since the available GAM implementation (Hastie, 2011) does not provide standard errors or confidence intervals for "new" objects that are not included in the training sample, we proceed as follows to estimate standard errors for each location in the prediction map. We first compute standard errors on the logit and probability levels for all sample points. A lookup table is then used to transfer these standard errors to all grid cells of the raster based on the similarity of the values of explanatory variables used by the model. Tolerance thresholds were applied to each explanatory variable to identify suitable observations in the training sample that match any given prediction location and therefore have similar standard errors. Several tolerance thresholds were tested for each modelling domain to maximize the R^2 obtained by comparing, on the training sample, standard errors estimated by table lookup to the standard errors calculated by the GAM implementation. This results in a raster data set which gives an estimation of the standard error of the predicted logit for each grid cell.

Based on these approximated standard errors we estimate the approximate upper and lower limit of a 95 % confidence interval of the predicted logit using a normal approximation. These logit-scale confidence intervals are further converted to the probability level and adjusted based on each modelling domain's sampling rate (Sect. 5.2). These approximate upper and lower confidence limits and the predicted probability are used to visualize and compare the spatially varying uncertainties in a classified landslide susceptibility map. The classified susceptibility map is compared to the classified maps of upper and lower confidence limits to assess potential areas and grid cells in which misclassification of the susceptibility class may occur.

Furthermore, the approximate logit-scale standard errors from each model's predictions are used as relative uncertainty indices of the susceptibility map within each domain. This uncertainty index allows for a more nuanced visualisation of prediction uncertainties within each domain, knowing, however, that its interpretation is only applicable within the domain since no adjustment for the domain-specific sampling rate is applied to this index.

6 Results

6.1 Landslide susceptibility map

The three classes of the final landslide susceptibility map classified according to the proportion of landslides included covered 75 % (low susceptibility), 19 % (medium susceptibility) and 6 % (high susceptibility) of the total study area (Fig. 3).

The variable frequency analysis showed that different subsets of the available 15 explanatory variables were included in the GAMs (GAM_1–GAM_{16}) for the different modelling domains (Table 2). The total number of variables used in the models (GAM_1–GAM_{16}) ranged from four variables in the Bohemian Massif with plutonic rock and Waschberg Zone including Bohemian Massif with sedimentary cover domains (52 landslides each) to 11 variables used in the model for the Flysch Zone (6281 landslides, Table 3). The number of variables included in the models generally increased with the number of observations in the training sample, which was attributed to the AIC's penalization based on model complexity relative to sample size. Furthermore, 65 % of the variables were used in a smoothed form (Table 2). However, mainly linear model terms were selected in four modelling domains. A similar overall frequency of nonlinear model terms was obtained in the models fitted within the cross-validation procedures (71 % nonlinear overall) with very similar results for the spatial and non-spatial techniques (refer to Table 2 for details). Two domains (Loess, Loam and Waschberg Zone and Bohemian Massif with sedimentary cover) primarily used linear model terms in spatial and non-spatial cross-validation. Additionally, in the Bohemian Massif with plutonic rock (only in spatial cross-validation) and the alluvial deposits including lake and

NHESSD
1, 1001–1050, 2013

Assessing the quality of landslide susceptibility maps of Lower Austria

H. Petschko et al.

wetland (only in non-spatial cross-validation) high proportions of linear terms were observed.

6.2 Spatial and non-spatial cross-validation

In general, spatial cross-validation had a larger range of AUROC values than non-spatial cross-validation over all cross-validation repetitions (Fig. 4). However, the median AUROC values were very similar; non-spatial cross-validation had only slightly higher median values. The highest median AUROC values of 0.98 (spatial) and 0.99 (non-spatial cross-validation) were found in the alluvial deposits including lakes and wetland domain. With the exception of the Permo-Mesozoic rocks domain, all median AUROC values are higher than 0.74 (Fig. 4). In this domain the largest differences of median AUROC values between spatial (0.53) and non-spatial cross-validation (0.79) were found. Furthermore, the lowest 1st quartile AUROC value, which was higher performing with non-spatial cross-validation (0.73) than spatial cross-validation (0.35), was found for the Permo-Mesozoic rocks domain. Additionally, this domain showed the highest interquartile range of AUROC values (0.42) with spatial cross-validation and the second highest (0.11) with non-spatial cross-validation.

Three domains had very high AUROC values of the 3rd quartile, which ranged from 0.97 to 1 in the spatial and non-spatial cross-validation. These were the loess and loam, the Bohemian Massif with plutonic rock and the alluvial deposits including lakes and wetland domains.

The poorest spatial and non-spatial transferabilities assigned at $T_{sp} > 0.10$ and $T_{nsp} > 0.04$ were obtained in the three modelling domains with the smallest sample sizes (Table 3). Transferability tended to be better in domains with larger sample sizes and/or higher landslide densities, but was unrelated to median AUROC. Among the domains with larger sample sizes, the Austroalpine Unit with dolostone stood out with relatively poor spatial transferability ($T_{sp} = 0.098$) in addition to its relatively low median AUROC of 0.75. Spatial transferabilities were best ($T_{sp} < 0.03$) for igneous rocks of the Austroalpine Unit, the Molasse zone and the Schlier Zone.

Reducing the sample size from using the total number of landslides in the case-control sample (12562) to 50 samples within a modelling domain (Flysch Zone) still resulted in acceptable median AUROC values (> 0.76) for all sample sizes. However, the median AUROC values decreased from 0.84 to 0.76 in both the spatial and non-spatial cross-validation (Fig. 5) with very little difference between both approaches. While the AUROC values stay relatively constant until a sample size of 3200, they started to decrease more rapidly from there on as the sample size decreases. Despite that, the interquartile ranges were substantially higher with spatial cross-validation, and generally increased with decreasing sample sizes in both spatial (0.052–0.087) and non-spatial (0.011–0.059) cross-validation. Below a sample size of about 400 (spatial cross-validation) and 200 (non-spatial cross-validation) the interquartile range of the AUROC values sharply increases as the sample size decreased; thus the transferability of the model decreased substantially for small sample sizes. In addition, the smaller sample sizes led to lower thematic consistency of the model.

6.3 Variable importance and thematic consistency

In the comparison of the models fitted within the 16 modelling domains (GAM_1–GAM_{16}), topographic variables were more important than the available variables on soil properties. Out of a maximum of 16 selections, the variable slope angle was selected 15 times within stepwise variable selection, whereas the minimum permeability value was not selected at all (Table 2). However, this only shows the result of one specific random sample and variable selection repetition. According to the relative variable-selection frequency resulting from the two cross-validation approaches, the variable importance for predicting the landslide susceptibility also changed distinctly between the modelling domains. All modelling domains (except the domain of the Permo-Mesozoic rocks) had slope angle as the most important variable. It was selected on average in 91.8 % and 95.7 % of the model repetitions in spatial and non-spatial cross-validation. Other important variables were catchment height, which was selected 55.6 % (spatial) and 65 % (non-spatial) of the repetitions. In the spatial cross-validation the Euclidian

NHESSD
1, 1001–1050, 2013

distance to nappe boundaries was the second most important variable as it was selected in 74.2 % of the models (non-spatial selection frequency 56.1 %, rank 6). The topographic wetness index (68.3 %/61.4 %), the convergence index (50 pixel radius; 56.1 %/64.3 %) and the curvature (53.8 %/58.6 %) were also among the top ranking variables in both cross-validation approaches. Void space (mean 0–100 cm) was not selected in any of the model runs, while void space (0–20 cm) was selected by less than 1 % of the models and the minimum permeability was included in 1.1 % (spatial cross-validation) to 4.9 % (non-spatial cross-validation) of the models on average.

Overall the thematic consistency was stronger within the non-spatial cross-validation because training sets are less variable when the data are partitioned randomly as opposed to spatially (Table 3). A strong thematic consistency was assigned for a consistency index of C_{sp}, $C_{nsp} < 0.20$ and was found for seven domains in the non-spatial cross-validation but only for four domains in the spatial cross-validation. The strongest thematic consistency was observed in the Flysch Zone ($C_{sp} = 0.099$) in spatial cross-validation and in the Austroalpine Unit with dolostone ($C_{nsp} = 0.068$) and the Klippen Zone ($C_{nsp} = 0.081$) in the non-spatial cross-validation. These domains had a very large sample size and landslide density and also a good spatial or non-spatial transferability. While in spatial cross-validation the domains with the smallest sample size and poorest spatial transferability had the weakest thematic consistency, in non-spatial cross-validation the consistency index was unrelated to sample size and transferability. Among these the Waschberg Zone including the Bohemian Massif with sedimentary cover domain gave very contrasting results. It showed a weak thematic consistency in spatial cross-validation but a medium consistency index in non-spatial cross-validation. The weakest thematic consistency was found for the Permo-Mesozoic rocks ($C_{sp} = 0.519$) and the alluvial deposits including lakes and wetlands ($C_{nsp} = 0.389$) domains. In non-spatial cross-validation the thematic consistency was stronger with lower median AUROC values. However, in spatial cross-validation the median AUROC values and the thematic consistency were unrelated.

NHESSD
1, 1001–1050, 2013

6.4 Spatially varying prediction uncertainties

The largest range of logit-scale standard errors was obtained for the Bohemian Massif with plutonic rock domain (0.37–16.24), whereas the Quaternary fluvial terrace had the minimum range (0.22–1.69). The highest standard error of 16.24 was in the Bohemian Massif with plutonic rock domain, while the median of the highest standard errors over all domains was 2.27. After the transformation from logit-scale to probability scale the typical distribution of the standard error was the following: the range of the standard error of the predicted probability was largest for medium probabilities that were in the medium susceptibility class. The minimum range of the standard errors was typically obtained at the minimum and maximum probability that were contained in the low and high susceptibility class. Nevertheless, the lowest standard errors were computed at the minimum predicted probability.

The results of the analysis of spatially varying uncertainties are presented by maps chosen in two exemplary modelling domains with contrasting landslide densities, the Flysch Zone ($4.6\,\mathrm{km}^{-2}$) and the Bohemian Massif ($0.09\,\mathrm{km}^{-2}$, Fig. 6). The range of the standard errors of the predicted logit of these domains was very similar. While the standard errors range from 0.05 to 1.6 in the Flysch Zone, the Bohemian Massif had a range from 0.24 to 1.73. With this minimum value the Flysch Zone gives the lowest standard error in the study area.

Within the study area of Lower Austria seven types of susceptibility class uncertainties in the landslide susceptibility maps were identified by the analysis of overlaps between the susceptibility classes (Figs. 6 and 7). Most commonly, for about 85 % of the grid cells, there were no overlaps of different susceptibility classes either with the lower or the upper confidence limit meaning that in all maps the susceptibility class was the same. The most common overlaps were found for the low susceptibility class where 6 % of the grid cells were classified with low susceptibility in the predicted probability map but with medium susceptibility in the upper confidence limit map (Fig. 7). Moreover, 2 % of the study area experienced overlaps of the medium class in the predicted

NHESSD
1, 1001–1050, 2013

probability map to the high susceptibility class in the upper confidence limit map. Even less grid cells, 0.03 % had a range from the low class in the predicted probability map to the high class in the upper confidence limit map.

Comparing the maps from the Flysch Zone to the Bohemian Massif it can be seen that the grid cells showing an overlap in the Flysch Zone were much more scattered, pixel wise, and mainly occurred at the boundaries of the susceptibility classes. In the centre of the high susceptible class no overlaps occurred. However, in the Bohemian Massif the overlaps mainly occurred between the medium and high susceptibility classes and generally covered the entire class or larger areas, instead of single grid cells only. Both maps shared a low frequency of dark blue and dark green grid cells, which showed overlaps from the low class to the high class (either from lower confidence limit to predicted probability or from predicted probability to upper confidence limit). Furthermore, the overlaps of classes were occurring equally in- and outside of permanent settlement areas, which was important considering the map for future planning purposes (Fig. 6).

7 Discussion

7.1 Quality of the input data

Generally a complete, unbiased inventory would be desirable, as for example a full inventory that was mapped directly in the aftermath of a landslide event (single landslide or multiple landslides triggered at the same time) in the area of the susceptibility map. A complete inventory is rarely available. Particularly for historical inventories the level and type of completeness is unknown while it is known that they are generally incomplete (Malamud et al., 2004). Even a substantially complete inventory, which would be useful in statistical modelling as it includes a substantial fraction (random sample) of all landslides at all scales, land use types, lithological units or slope angles, cannot be reached for historical inventories (Malamud et al., 2004). This origins from the observation that landslides and their visibility on aerial photographs or other base maps (e.g. hillshades derived from airborne laser scanning DTMs) are highly influenced by new landslides, erosion, land use type and anthropogenic activities (Bell et al., 2012; Malamud et al., 2004; McCalpin, 1984). Furthermore the mapping and identification of landslides is highly dependent on the experience and knowledge of the investigator (Ardizzone et al., 2002). If these influences on the completeness are not random they might introduce a bias in the inventory and furthermore in the sampling which results in a biased or systematic modelling error. In our study area it is assumed, that the inventory is not complete, as it originates from recent data sources (not multi-temporal) only and the visibility of landslides in the ALS-DTM or orthophoto is influenced by human impact depending on the land use type (Bell et al., 2012). However, the type of incompleteness was not analysed for the entire study area. Therefore, it is not clear if the missing landslides are missing completely at random or are biased toward absence in certain land uses or lithological units.

The effect of a reduced sample size on the median and interquartile range AUROC values was assessed. We found that the median AUROC remained satisfactory high but decreased as sample size decreased, while the interquartile range of the AUROC increased. According to the median AUROC we found that even with the smallest sample size the model still achieved a good discrimination between landslide and non-landslide points. Furthermore, the AUROC remained constant decreasing the sample size by a quarter of the total sample size (3200). This might be related to the size of the test sample, which was set with 2000. However, the higher interquartile range AUROC values associated with lower sample sizes demonstrated that the spatial and non-spatial transferability and the thematic consistency decreased with shrinking sample size.

NHESSD
1, 1001–1050, 2013

Assessing the quality of landslide susceptibility maps of Lower Austria

H. Petschko et al.

7.2 Quality of the statistical model

7.2.1 Study design to meet the heterogeneity of the study area

Observed variable-selection frequencies showed that different explanatory variables were relevant in different domains, which provides additional justification to the decision to model susceptibility in each modelling domain separately. Additionally, not only the different choice of the variables is important but also the way the variables are fitted or smoothed according to the sample in the respective domain. Previous studies showed that within each lithological unit landslides occur at different slope angles (Muenchow et al., 2012; Petschko et al., 2012). Similar differences within lithological units or terrain types might be present for other explanatory variables as well, as the geomorphic and geologic characteristics change (Lee et al., 2008). Facing this, our study design gives much more flexibility to represent the characteristics of the study area.

However, one may argue that with this approach problems occur at the boundaries of the lithological units. Inaccuracies in the delineation of the lithological map of the area have effects on the model results as the landslides might be assigned incorrect lithological information. This may lead to an underestimation of effect sizes as data from different (true) lithological units would be mixed. Similar mixing effects may occur for quantitative predictor variables as well, for example as a function of positional accuracy for scale and resolution. In regression this effect is known as dilution, which may introduce a bias of estimated regression coefficients toward zero (Frost and Thompson, 2000). As this would also occur using the lithological map as a factor instead of as a mask for partitioning the study area, the boundary inaccuracies are not only a problem in the applied study design.

7.2.2 Spatial and non-spatial cross-validation

Cross-validation estimation of a model's predictive performance is a crucial step in predictive modelling because estimation on the training set is always too optimistic

NHESSD
1, 1001–1050, 2013

(Hosmer and Lemeshow, 2000; Brenning, 2005). Cross-validation results in bias-reduced performance estimates as the test partitions used in each repetition do not overlap with the training sample (Brenning, 2005). In particular spatial cross-validation is recommendable for spatial data, which may be subject to spatial autocorrelation (Brenning, 2005, 2012a).

While median AUROC values estimated by spatial and non-spatial cross-validation were very similar in this study, non-spatial cross-validation provided a more optimistic assessment of model transferability in contrast to spatial cross-validation. Therefore, spatial and temporal cross-validation should be preferred for performance estimation (Chung and Fabbri, 2008). While spatial performance and transferability do not necessarily reflect a model's predictive performance in the time domain, they provide a more realistic assessment of its ability to generalize from the available data than non-spatial approaches (Brenning, 2005).

The spatially and non-spatially least transferable models in this study were associated with domains that had the smallest sample sizes. The relationship of sample size on predictive abilities has also been shown in other spatial modelling studies (Stockwell and Peterson, 2002; Hjort and Marmion, 2008). However, we believe that the cases of high variation in AUROC values may be also related to the cross-validation sampling variation as indicated by the difference between T_{sp} and lower T_{nsp}, and possibly the proportion of stable and unstable terrain in a modelling domain. Heterogeneity of landslide conditions (e.g. related to topography or land use) in the cross-validation samples is more likely to occur if samples are partitioned spatially, such as the case in spatial cross-validation. This heterogeneity between the training and test sample partitions may have adverse effects on the derivation of the performance estimator (Guzzetti et al., 2006). The model domains that have high contrast between stable (e.g. large flat areas) and unstable (e.g. steep areas) terrain have potential for greater variation of sampled terrain conditions; it may be possible that in some samples one terrain condition is overrepresented relative to others. The sampling strategy may be improved by masking low-lying flat areas that are not typically susceptible to landslides (Van den

NHESSD
1, 1001–1050, 2013

Eeckhaut et al., 2009; Goetz et al., 2011). Consequently, the sample may have more potential to capture the differentiating conditions of stable and unstable terrain in an area that is generally susceptible to landslides (e.g. steep hillslopes). Since each of the k folds is used once as a test sample, this estimation procedure effectively uses the entire data set for testing and for training, which is an advantage over traditional approaches (i.e. Van den Eeckhaut et al., 2007a; Frattini et al., 2010; Rossi et al., 2010) using a fixed test set (Brenning, 2005). We therefore introduced a transferability index based on the interquartile range of the performance estimator (AUROC) to assess the transferability of the models under a variety of sampling conditions.

A similar relation was found between the thematic consistency and the sampling size and landslide density. Whereas a very strong thematic consistency was found for domains with a large sample size and sampling rate, domains with small sample size and rate showed a high variability in the variable-selection frequencies which gave a weak thematic consistency. Therefore, the weak thematic consistency might also be associated with a poor spatial and non-spatial transferability, both originating from a small sample size and a small sampling rate. This relation was stronger for the spatial cross-validation, while the thematic consistency from non-spatial cross-validation was unrelated. The thematic consistency in non-spatial cross-validation tended to be stronger with a lower median AUROC value.

7.3 Quality of the map – susceptibility class uncertainty

This analysis of susceptibility class uncertainty is an improvement of previous landslide susceptibility studies (e.g. Guzzetti et al., 2006; Van den Eeckhaut et al., 2009; Rossi et al., 2010; Sterlacchini et al., 2011), as it not only showed an uncertainty index of the predicted probability on a grid cell basis but additionally provided information on where the susceptibility class uncertainties were located. According to this new information it was found that in the classified map the majority of grid cells did not change. Furthermore, special attention has to be put in the low susceptibility class, as here an underestimation of the susceptibility might occur. This is of special interest for future

NHESSD
1, 1001–1050, 2013

land-use and development planning usually performed by non-landslide experts of local governments.

In the European context some examples of the implementation of landslide susceptibility maps for earth and debris slides but also for other natural hazards (debris flows and avalanches) in land-use planning can be found among others in Switzerland, in the Flemish Ardennes, in the Bavarian Alps, in South Tyrol or in Austria (mainly debris flows and avalanches) (Bundesamt für Raumentwicklung et al., 2005; Van den Eeckhaut et al., 2007b; Bayerisches Landesamt für Umwelt (LfU), 2012; Autonome Provinz Bozen – Südtirol, 2012; Rudolf-Miklau, 2007). In Switzerland as well as in South Tyrol so called "hazard zonation plans" are compulsory for the municipalities and outline areas where the designation of new building areas is restricted or only allowed with special adaptations (Bundesamt für Raumentwicklung et al., 2005; Autonome Provinz Bozen- Südtirol, 2012). In all examples uncertainties in the boundaries of the delineated susceptibility or hazard zones are not analysed in detail and only vaguely communicated. Interviews of Klimeš and Blahůt (2012) showed that local governments even do not want this information. However, these uncertainties might have severe consequences on buildings and their inhabitants if an event occurred within the uncertainties of the method used to delineate the hazard zones.

Similarly to the listed examples, in the aftermath of this study each landslide susceptibility class will be related to, not legally binding, recommendations for the designation of new building areas. Therefore, a misclassification (e.g. low instead of medium susceptibility) might lead to an interpretation by the municipality or landowner that underestimated landslide susceptibility. Knowledge about the susceptibility class uncertainties might outline where more caution and detailed investigations are necessary. Additionally, it also shows where no uncertainties are expected, which might help to avoid costs for slope investigations. Furthermore, this analysis might aid to a good acceptance of the landslide susceptibility maps in the local governments, as instead of a fuzzy statement on involved uncertainties these are clearly shown in a map on grid cell level (Luoto et al., 2010).

NHESSD
1, 1001–1050, 2013

8 Conclusions

The quality of landslide susceptibility maps was analysed in terms of landslide susceptibility model performance and susceptibility class uncertainties of the final classified susceptibility map. The landslide susceptibility model and the resulting classified map of this study are regarded to be of high quality.

The applied study design with modelling in the different domains provides a high flexibility for representing the characteristics of the heterogeneous study area. The spatial cross-validation gave a more realistic assessment of the model performance and spatial transferability from the available data than non-spatial approaches. The results showed a weak relationship between sample sizes and sampling rates on the one hand and spatial transferability on the other. However, in spatial cross-validation a tendency of a stronger thematic consistency with larger sample sizes and sampling rates was found. Furthermore, reducing the sample size within a single modelling domain resulted in lower but still good median AUROC value and thematic consistency but a larger interquartile range of AUROC values which gave a lower spatial and non-spatial transferability. Regarding the susceptibility class uncertainties we conclude that the majority of the study area is not affected by class uncertainties. Special attention has to be drawn to possible overlaps of the low and medium susceptibility class in the predicted probability map and the map of the upper confidence limit. A misclassification in the low class might result in an underestimation of the susceptibility. This might have adverse effects on the municipality or landowner if the recommendations for the assignment of building areas might not be restrictive enough. Therefore, detailed knowledge on inherent susceptibility class uncertainties within one modelling method as presented and visualized in this study is of high importance for well-informed future planning and decision making.

Acknowledgements. This study was conducted within the research project MoNOE – Method development for landslide susceptibility maps for Lower Austria funded by the Provincial Government of Lower Austria. The data on the topography, soil and tectonic lines was by courtesy of the Provincial Government of Lower Austria. Furthermore, the geological data was provided by the Geological Survey of Austria (GBA-2009 – Zl.383/1-09). The authors are thankful for the support and provision of additional data (landslide archive, simplified geological map and land cover map) of the Geological Survey of Lower Austria and the Department of Spatial Planning and Regional Policy and of the project partners Austrian Institute of Technology and Joanneum Research.

References

Abteilung Feuerwehr und Zivilschutz, Amt der NÖ Landesregierung: Zusammenfassung Ereignisse 2009, available at: http://www.noe.gv.at/Land-Zukunft/Katastrophenschutz/Archiv/zusammenfassung_ereignisse_2009.wai.html (last access: 9 December 2010), 2010 (in German).

Akaike, H.: A new look at the statistical model identification, IEEE T. Automat. Contr., 19, 716–723, 1974.

Ardizzone, F., Cardinali, M., Carrara, A., Guzzetti, F., and Reichenbach, P.: Impact of mapping errors on the reliability of landslide hazard maps, Nat. Hazards Earth Syst. Sci., 2, 3–14, doi:10.5194/nhess-2-3-2002, 2002.

Autonome Provinz Bozen – Südtirol: Beschluss der Landesregierung Autonome Provinz Bozen – Südtirol: Abänderung der Richtlinien zur Erstellung der Gefahrenzonenpläne gemäß Landesraumordnungsgesetz, Landesgesetz vom 11. August 1997, Nr. 13, Artikel 22/bis, Beiblatt Nr. 1, Autonome Provinz Bozen – Südtirol, Südtirol, Italien, 2012 (in German).

Bayerisches Landesamt für Umwelt (LfU) (Ed.): Gefahrenhinweiskarte Bayerische Alpen – Schutz vor Georisiken: Erfahrungen und praktische Umsetzung, Bayerisches Landesamt für Umwelt (LfU), Augsburg, Deutschland, 2012 (in German).

Beguería, S.: Validation and evaluation of predictive models in hazard assessment and risk management, Nat. Hazards, 37, 315–329, 2006.

Bell, R.: Lokale und regionale Gefahren- und Risikoanalyse gravitativer Massenbewegungen an der Schwäbischen Alb, Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn, available at: http://hss.ulb.uni-bonn.de/2007/1107/1107.pdf, last access: 20 March 2012, 2007 (in German).

NHESSD
1, 1001–1050, 2013

Bell, R., Petschko, H., Röhrs, M., and Dix, A.: Assessment of landslide age, landslide persistence and human impact using airborne laser scanning digital terrain models, Geogr. Ann. A, 94, 135–156, 2012.

BMLFUW: Nachhaltig geschützt – Naturgefahrenmanagement im Unwetterjahr 2009, Jahresbericht, Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft (Lebensministerium), Wien, available at: http://www.forstnet.at/article/articleview/82437/1/4932 (last access: 3 December 2010), 2010 (in German).

Boehner, J., Koethe, R., Conrad, O., Gross, J., Ringeler, A., and Selige, T.: Soil Regionalisation by Means of Terrain Analysis and Process Parameterisation, Research Report, European Soil Bureau, Luxembourg, 2002.

Brabb, E. E.: Innovative approaches to landslide hazard mapping, 4th International Symposium on Landslides, 16–21 September , Toronto, Canada, 307–324, 1984.

Brenning, A.: Spatial prediction models for landslide hazards: review, comparison and evaluation, Nat. Hazards Earth Syst. Sci., 5, 853–862, doi:10.5194/nhess-5-853-2005, 2005.

Brenning, A.: Statistical Geocomputing combining R and SAGA: The Example of Landslide susceptibility Analysis with generalized additive Models, SAGA – Seconds Out, Hamburger Beiträge zur Physischen Geographie und Landschaftsökologie, 19, 23–32, 2008.

Brenning, A.: Benchmarking classifiers to optimally integrate terrain analysis and multispectral remote sensing in automatic rock glacier detection, Remote Sens. Environ., 113, 239–247, 2009.

Brenning, A.: RSAGA: SAGA Geoprocessing and Terrain Analysis in R, Version 0.93-1, available at: http://CRAN.R-project.org/package=RSAGA, last access: 22 January 2013, 2011.

Brenning, A.: Improved spatial analysis and prediction of landslide susceptibility: practical recommendations, in: Landslides and Engineered Slopes, Protecting Society through Improved Understanding, edited by: Eberhardt, E., Froese, C., Turner, A. K., and Leroueil, S., Taylor & Francis, Banff, Alberta, Canada, 789–795, 2012a.

Brenning, A.: Spatial cross-validation and bootstrap for the assessment of prediction rules in remote sensing: the R package "sperrorest", IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 23–27 July 2012, Munich, 5372–5375, 2012b.

Brierley, G.: Communicating Geomorphology, J. Geogr. Higher Educ., 33, 3–17, 2009.

Bundesamt für Raumentwicklung, Bundesamt für Wasser und Geologie, and Bundesamt für Umwelt, Wald und Landschaft (Eds.): Empfehlung Raumplanung und Naturgefahren, BBL, Verkauf Bundespublikationen, Bern, Switzerland, 2005 (in German).

Carrara, A., Cardinali, M., Guzzetti, F., and Reichenbach, P.: GIS technology in mapping landslide hazard, in: Geographical Information Systems in Assessing Natural Hazards, edited by: Carrara, A. and Guzzetti, F., Kluwer Academic Publishers, Dordrecht, Netherlands, 135–175, 1995.
Carrara, A., Guzzetti, F., Cardinali, M., and Reichenbach, P.: Use of GIS technology in the prediction and monitoring of landslide hazard, Nat. Hazards, 20, 117–135, 1999.
Chung, C. J. F. and Fabbri, A. G.: Probabilistic prediction models for landslide hazard mapping, Photogramm. Eng. Rem. S., 65, 1389–1399, 1999.
Chung, C. J. F. and Fabbri, A. G.: Validation of spatial prediction models for landslide hazard mapping, Nat. Hazards, 30, 451–472, 2003.
Chung, C. J. and Fabbri, A. G.: Predicting landslides for risk analysis–spatial models tested by a cross-validation technique, Geomorphology, 94, 438–452, 2008.
Conrad O.: SAGA – Entwurf, Funktionsumfang und Anwendung eines Systems für Automatisierte Geowissenschaftliche Analysen, Ph.D. thesis, University of Göttingen, Germany, available at: http://hdl.handle.net/11858/00-1735-0000-0006-B26C-6, last access: 27 February 2013, 2007 (in German).
Crozier, M. J.: Landslides: causes, consequences and environment, Croom Helm, London, Sydney, Dover, New Hampshire, 1986.
Cruden, D. M. and Varnes, D. J.: Landslide types and processes, in: Landslides, Investigation and Mitigation, edited by: Turner, A. K. and Schuster, R. L., Transportation Research Board Special Report 247, Washington D.C., USA, 36–75, 1996.
Dai, F. C., Lee, C. F., and Ngai, Y. Y.: Landslide risk assessment and management: an overview, Eng. Geol., 64, 65–87, 2002.
Damm, A., Eberhard, K., and Patt, A.: Risikowahrnehmung von Erdrutschen: Ergebnisse einer empirischen Untersuchung in der Südoststeiermark, Wegener Zentrum für Klima und Globalen Wandel, International Institute for Applied Systems Analysis (IIASA), Graz, Wien, 2010 (in German).
Demoulin, A. and Chung, C.-J. F.: Mapping landslide susceptibility from small datasets: a case study in the Pays de Herve (E Belgium), Geomorphology, 89, 391–404, 2007.
Dikau, R., Brunsden, D., Schrott, L., and Ibsen, M. L. (Eds.): Landslide Recognition: Identification, Movement, and Clauses, Wiley, Chichester, 1996.
Eder, A., Sotier, B., Klebiner, K., Sturmlechner, R., Dorner, J., Markat, G., Schmid, G., and Strauss, P.: Hydrologische Bodenkenndaten der Böden Niederösterreichs (HydroBodNÖ)

NHESSD
1, 1001–1050, 2013

Assessing the quality of landslide susceptibility maps of Lower Austria

H. Petschko et al.

(Data on hydrological soil characteristics of soils in Lower Austria), Final Report, Bundesamt für Wasserwirtschaft, Institut für Kulturtechnik und Bodenwasserhaushalt, Bundesforschungszentrum für Wald, Institut für Naturgefahren, Petzenkirchen, Innsbruck, 2011 (in German).
Egner, H. and Pott, A.: Risiko und Raum: Das Angebot einer Beobachtungstheorie, in: Geographische Risikoforschung – Zur Konstruktion verräumlichter Risiken und Sicherheiten, edited by: Egner, H. and Pott, A., Franz Steiner Verlag, Stuttgart, 9–35, 2010 (in German).
Fabbri, A. G., Chung, C. J. F., Cendrero, A., and Remondo, J.: Is prediction of future landslides possible with a GIS?, Nat. Hazards, 30, 487–503, 2003.
Frattini, P., Crosta, G., and Carrara, A.: Techniques for evaluating the performance of landslide susceptibility models, Eng. Geol., 111, 62–72, 2010.
Freeman, G. T.: Calculating catchment area with divergent flow based on a regular grid, Comput. Geosci., 17, 413–422, 1991.
Frost, C. and Thompson, S.: Correcting for regression dilution bias: comparison of methods for a single predictor variable, J. R. Stat. Soc. Ser. A-G., 163, 173–190, 2000.
Glade, T. and Crozier, M. J.: A review of scale dependency in landslide hazard and risk analysis, in: Landslide Hazard and Risk, edited by: Glade, T., Anderson, M., and Crozier, M. J., John Wiley & Sons, Ltd, Chichester, England, 75–138, 2005.
Glade, T., Anderson, M. G., and Crozier, M. J. (Eds.): Landslide Hazard and Risk, John Wiley & Sons, Ltd, Chichester, England, 2005.
Glade, T., Petschko, H., Bell, R., Bauer, C., Granica, K., Heiss, G., Leopold, P., Pomaroli, G., Proske, H., and Schweigl, J.: Landslide susceptibility maps for Lower Austria – methods and challenges, in: 12th Congress Interpraevent 2012 Conference Proceedings, vol. 1, edited by: Koboltschnig, G., Hübl, J., and Braun, J., International Research Society INTERPRAEVENT, Grenoble, France, 497–508, 2012.
Goetz, J. N., Guthrie, R. H., and Brenning, A.: Integrating physical and empirical landslide susceptibility models using generalized additive models, Geomorphology, 129, 376–386, 2011.
Gottschling, P.: Massenbewegungen, in: Geologie der Bundesländer – Niederösterreich, Geologische Bundesanstalt, Wien, 335–340, 2006 (in German).
Guzzetti, F.: Landslide hazard and risk assessment, Dissertation, Rheinischen Friedrich-Wilhelms-Universität Bonn, Bonn, 2005.

NHESSD
1, 1001–1050, 2013

Guzzetti, F., Carrara, A., Cardinali, M., and Reichenbach, P.: Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy, Geomorphology, 31, 181–216, 1999.
Guzzetti, F., Cardinali, M., Reichenbach, P., and Carrara, A.: Comparing landslide maps: a case study in the upper Tiber River Basin, Central Italy, Environ. Manage., 25, 247–263, 2000.
Guzzetti, F., Reichenbach, P., Ardizzone, F., Cardinali, M., and Galli, M.: Estimating the quality of landslide susceptibility models, Geomorphology, 81, 166–184, 2006.
Hand, D.: Construction and Assessment of Classification Rules, Wiley, West Sussex, 1997.
Hanley, J. A. and McNeil, B. J.: The meaning and use of the area under a receiver operating characteristic (ROC) curve, Radiology, 143, 29–36, 1982.
Hastie, T.: gam: Generalized Additive Models, available at: http://CRAN.R-project.org/package=gam, last access: 27 February 2013, 2011.
Hastie, T. and Tibshirani, R.: Generalized Additive Models, Chapman & Hall/CRC, London, 1990.
Hervás, J.: Lessons Learnt from Landslide Disasters in Europe, European Commission Joint Research Centre, Institute for Environment and Sustainability, Ispra (VA), Italy, 2003.
Hjort, J. and Marmion, M.: Effects of sample size on the accuracy of geomorphological models, Geomorphology, 102, 341–350, 2008.
Hornich, R. and Adelwöhrer, R.: Landslides in Styria in 2009/Hangrutschungsereignisse 2009 in der Steiermark, Geomechanics and Tunnelling, 3, 455–461, 2010 (in German).
Hosmer, D. W. and Lemeshow, S.: Applied Logistic Regression, Wiley, New York, NY, 2000.
Hydrographischer Dienst des Landes Niederösterreich: Wasserstandsnachrichten und Hochwasserprognosen Niederösterreich, available at: http://www.noel.gv.at/Externeseiten/wasserstand/wiskiwebpublic/maps_N_0.htm?entryparakey=N (last access: 2 March 2011), 2011 (in German).
Jia, G., Tian, Y., Liu, Y., and Zhang, Y.: A static and dynamic factors-coupled forecasting model of regional rainfall-induced landslides: a case study of Shenzhen, Sci. China Ser. E, 51, 164–175, 2008.
Klimeš, J. and Blahůt, J.: Landslide risk analysis and its application in regional planning: an example from the highlands of the Outer Western Carpathians, Czech Republic, Nat. Hazards, 64, 1779–1803, 2012.
Knuepfer, P. L. and Petersen, J. F.: Geomorphology in the public eye: policy issues, education, and the public, Geomorphology, 47, 95–105, 2002.

Kunz, M., Grêt-Regamey, A., and Hurni, L.: Visualization of uncertainty in natural hazards assessments using an interactive cartographic information system, Nat. Hazards, 59, 1735–1751, 2011.
Kurz, W.: Erstellung einer digitalen, strukturgeologischen, tektonischen Karte von Niederösterreich (Endbericht), unpublished final report, Karl-Franzens-University Graz, Graz, 2012 (in German).
Malamud, B. D., Turcotte, D. L., Guzzetti, F., and Reichenbach, P.: Landslide inventories and their statistical properties, Earth Surf. Proc. Land., 29, 687–711, 2004.
Lee, C.-T., Huang, C.-C., Lee, J.-F., Pan, K.-L., Lin, M.-L., and Dong, J.-J.: Statistical approach to earthquake-induced landslide susceptibility, Eng. Geol., 100, 43–58, 2008.
Luoto, M., Marmion, M., and Hjort, J.: Assessing spatial uncertainty in predictive geomorphological mapping: a multi-modelling approach, Comput. Geosci., 36, 355–361, 2010.
McCalpin, J.: Preliminary age classification of landslides for inventory mapping, in: Proceedings 21st annual Enginnering Geology and Soils Engineering Symposium, 5–6 April,University of Idaho, Moscow, 99–11, 1984.
Muenchow, J., Brenning, A., and Richter M.: Geomorphic process rates of landslides along a humidity gradient in the tropical Andes, Geomorphology, 139–140, 271–284, 2012.
Orme, A. R.: Shifting paradigms in geomorphology: the fate of research ideas in an educational context, Geomorphology, 47, 325–342, 2002.
Park, N.-W. and Chi, K.-H.: Quantitative assessment of landslide susceptibility using high-resolution remote sensing data and a generalized additive model, Int. J. Remote Sens., 29, 247–264, 2008.
Petschko, H., Glade, T., Bell, R., Schweigl, J., and Pomaroli, G.: Landslide inventories for regional early warning systems, in: Proceedings of the International Conference "Mountain Risks: Bringing Science to Society", Firenze, 24–26 November 2010, edited by: Malet, J. P., Glade, T., and Casagli, N., CERG Editions, Strasbourg, 277–282, 2010.
Petschko, H., Bell, R., Brenning, A., and Glade, T.: Landslide susceptibility modeling with generalized additive models – facing the heterogeneity of large regions, in: Landslides and Engineered Slopes, Protecting Society through Improved Understanding, vol. 1, edited by: Eberhardt, E., Froese, C., Turner, A. K., and Leroueil, S., Taylor & Francis, Banff, Alberta, Canada, 769–777, 2012.

NHESSD
1, 1001–1050, 2013

Assessing the quality of landslide susceptibility maps of Lower Austria

H. Petschko et al.

Petschko, H., Bell, R., Glade, T., Leopold, P., and Heiss, G.: Landslide inventories for large regions – efficiency of mapping on the basis of ALS DTMs tested in Lower Austria, in preparation, 2013.
Petschko, H., Bell, R., Leopold, P., Heiss, G., and Glade, T.: Landslide inventories for reliable susceptibility maps in Lower Austria, edited by: Catani, F., Margottini, C., Trigila, A., and Iadanza, C., Springer, in press, 2013.
Quinn, P. F., Beven, K. J., Chevallier, P., and Planchon, O.: The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models, Hydrol. Process., 5, 59–79, 1991.
R Development Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, available at: http://www.R-project.org, last access: 27 February 2013, 2011.
Razak, K. A., Straatsma, M. W., Van Westen, C. J., Malet, J.-P., and De Jong, S. M.: Airborne laser scanning of forested landslides characterization: terrain model quality and visualization, Geomorphology, 126, 186–200, 2011.
Remondo, J., González, A., De Terán, J. R. D., Cendrero, A., Fabbri, A., and Chung, C. J. F.: Validation of landslide susceptibility maps; examples and applications from a case study in Northern Spain, Nat. Hazards, 30, 437–449, 2003.
Rogers, K. H.: The real river management challenge: integrating scientists, stakeholders and service agencies, River Res. Appl., 22, 269–280, 2006.
Rossi, M., Guzzetti, F., Reichenbach, P., Mondini, A. C., and Peruccacci, S.: Optimal landslide susceptibility zonation based on multiple forecasts, Geomorphology, 114, 129–142, 2010.
Rudolf-Miklau, F.: Austrian Service for Torrent and Avalanche Control, Federal Ministry of Agriculture, Forestry, Environment and Water Management, Vienna, available at: www.die.wildbach.lebensministerium.at, last access: 9 March 2011, 2007.
Ruß, G. and Brenning, A.: Data mining in precision agriculture: management of spatial information, Lect. Notes Comput. Sc., 6178, 350–359, 2010.
Schnabel, W.: Die geologischen Ursachen der Schäden an der II. Wiener Hochquellenleitung bei Scheibbs, in: Festschrift der MA 31: 75 Jahre II. Wiener Hochquellenleitung, edited by: Magistrat der Stadt Wien, MA 31-Wasserwerke, Stadt Wien, Wien, 1985 (in German).
Schnabel, W.: Geologische Karte von Niederösterreich 1 : 200 000, Geologische Bundesanstalt, Wien, Austria, 2002 (in German).

NHESSD
1, 1001–1050, 2013

Schwab, J. C., Gori, P. L., and Jeer, S.: Landslide Hazards and Planning, Planning Advisory Service Report, American Planning Association, Chicago, 2005.
Schwarz, L. and Tilch, N.: Möglichkeiten und Limitierungen der Regionalisierung mittels Neuronaler Netze am Beispiel einer Rutschungsanfälligkeitskarte für die Region Gasen-Haslau, Angewandte Geoinformatik, Beiträge zum 20. AGIT-Symposium, Angewandte Geoinformatik 2–4 July 2008, Salzburg, Austria, 643–648, 2008 (in German).
Schweigl, J. and Hervás, J.: Landslide Mapping in Austria, JRC Scientific and Technical Reports, European Commission Joint Research Centre, Institute for Environment and Sustainability, Italy, available at: http://eusoils.jrc.ec.europa.eu/ESDB_Archive/eusoils_docs/other/EUR23785EN.pdf, last access: 1 March 2011, 2009.
Schwenk, H.: Massenbewegungen in Niederösterreich 1953–1990, in: Jahrbuch der Geologischen Bundesanstalt, vol. 135, Geologische Bundesanstalt, Wien, 597–660, 1992 (in German).
Scott, A. J. and Wild, C. J.: Fitting logistic models under case-control or choice based sampling, J. Roy. Stat. Soc. B, 48, 170–182, 1986.
Sterlacchini, S., Ballabio, C., Blahut, J., Masetti, M., and Sorichetta, A.: Spatial agreement of predicted patterns in landslide susceptibility maps, Geomorphology, 125, 51–61, 2011.
Stockwell, D. R. and Peterson, A. T.: Effects of sample size on accuracy of species distribution models, Ecol. Model., 148, 1–13, 2002.
Van den Eeckhaut, M., Vanwalleghem, T., Poesen, J., Govers, G., Verstraeten, G., and Vandekerckhove, L.: Prediction of landslide susceptibility using rare events logistic regression: a case-study in the Flemish Ardennes (Belgium), Geomorphology, 76, 392–410, 2006.
Van den Eeckhaut, M., Poesen, J., Verstraeten, G., Vanacker, V., Nyssen, J., Moeyersons, J., Van Beek, L. P. H., and Vandekerckhove, L.: Use of LIDAR-derived images for mapping old landslides under forest, Earth Surf. Proc. Land., 32, 754–769, 2007a.
Van den Eeckhaut, M., Poesen, J., and Verstraeten, G.: Opstellen van een gevoeligheidskaart met betrekking tot massabewegingen (masstransport) voor de Vlaamse Ardennen, Departement Leefmilieu, Natuur en Energie, Afdeling Land en Bodembescherming, Ondergrond, Natuurlijke Rijkdommen, available at: http://www.lne.be/themas/bodem/grondverschuiving/grondverschuiving, last access: 9 March 2011, 2007b (in Dutch).
Van den Eeckhaut, M., Moeyersons, J., Nyssen, J., Abraha, A., Poesen, J., Haile, M., and Deckers, J.: Spatial patterns of old, deep-seated landslides: a case-study in the northern Ethiopian highlands, Geomorphology, 105, 239–252, 2009.

NHESSD
1, 1001–1050, 2013

Varnes, D. J.: Landslilde hazard zonation: a review of principles and practice, United Nations Educational, Scientific and Cultural Organization, Paris, France, 1984.
Von Ruette, J., Papritz, A., Lehmann, P., Rickli, C., and Or, D.: Spatial statistical modeling of shallow landslides – validating predictions for different landslide inventories and rainfall events, Geomorphology, 133, 11–22, 2011.
Vorpahl, P., Elsenbeer, H., Märker, M., and Schröder, B.: How can statistical models help to determine driving factors of landslides?, Ecol. Model., 239, 27–39, 2012.
Wessely, G.: Geologie der österreichischen Bundesländer – Niederösterreich, Geologische Bundesanstalt, Wien, 2006.
Zevenbergen, L. W. and Thornes, J. B.: Quantitative analysis of land surface topography, Earth Surf. Proc. Land., 12, 47–56, 1987.

NHESSD
1, 1001–1050, 2013

Assessing the quality of landslide susceptibility maps of Lower Austria

H. Petschko et al.

Table 1. Lithological units of Lower Austria: landslides and topography.

Code	Merged units	Name	Material*	Area (km^2)	No. of slides	Landslide density ($No. km^{-2}$)	Median slope angle (°)
1	0124	Anthropogenic deposits		2.93	0	0	3.1
2	02297	Alluvial deposits	Gravel, sand	3739.17	194	0.05	0.9
10	10	Loess, Loam	Sand, loam, loess-loam, drifting sand	2849.22	329	0.12	2.8
24	0124	Quaternary fluvial terrace	Gravel, sand, loess loam	747.79	222	0.30	2.2
35	35	Debris, till	Debris, till, scree material, rock avalanche material	217.19	177	0.81	16.6
37	3786	Bohemian Massif with sedimentary cover	Sandstone, claystone, conglomerate	30.95	0	0	12.2
39	39	Molasse Zone	Marl, sand, gravel, silt	1462.46	428	0.29	4.7
58	58	Molasse, Schlier	Fine sandy-silty marl in the circum-Alpine Tertiary basins	117.43	501	4.27	4.8
86	3786	Waschberg Zone	Marl, sand, limestone, clay	123.50	52	0.42	5.8
104	104	Intramontane Basins	Sand, gravel, breccia, clay, marl	737.26	291	0.39	5.2
120	120	Mélange Zone, Klippen Zone	Dominantly penninic metasediments and phiolites as well as insignificant Austroalpine elements, Ybbsitzer and Grestener Klippen Zone	73.46	404	5.50	11.9
126	126	Rheno-danubian Flysch Zone	Interbedded sandstone, marlstone to mudstone, marl	1365.96	6281	4.60	12.6
179	179	Austroalpine Unit with limestone, marls and sandstone	Limestone, marl, shale, sandstone, gypsum, conglomerate	785.85	1636	2.08	20.2
191	191	Austroalpine Unit with dolostone	Dolostone, limestone	2148.57	1419	0.66	27.1
230	230	Permo-Mesozoic rocks (overlying the Austroalpine ingeous rocks)	Carbonate Rocks, siliciclastics, porphyry (mostly metamorphics)	116.43	88	0.76	19.5
239	239	Igneous rocks of the Austroalpine Unit	Orthogneiss, Paragneiss, Mica-schist, Phyllite	614.35	586	0.95	15.1
251	251259	Bohemian Massif, Fault Zone	Tectonic fault zone	11.81	0	0	7.0
259	251259	Bohemian Massif	Paragneiss, mica-schist, phyllite, orthogneiss, Gföhl Gneiss, Granulite	2398.93	227	0.09	6.1
276	276	Bohemian Massif, plutonic rock	Granite, plutonic rock	1606.85	52	0.03	6.5
297	02297	Lake, wetland	Limnic sediments, wetland	38.77	2	0.05	3.5

* Sorted according to frequency of occurrence, after geological map.

Table 2. Variable frequency for the model using all landslide points (GAM_1–GAM_{16}) and variable-selection frequency of variables used linearly (*N*) or with a smoothing function (*S*) in spatial (spCV) and non-spatial (nspCV) cross-validation. All values are summarized over all modelling domains.

	Variable frequency (GAM_1–GAM_{16})		Variable frequency total (GAM_1–GAM_{16})	Relative variable-selection frequency from cross-validation					
	N	S		All spCV	All nspCV	N spCV	N nspCV	S spCV	S nspCV
Slope angle	2	13	15	91.8	95.7	9.1	17.3	82.8	78.4
Curvature	2	10	12	53.8	58.6	7.1	1.5	46.7	57.1
Topographic wetness index	4	7	11	68.3	61.4	15.9	13.5	52.3	47.9
Catchment height	4	6	10	55.6	65.0	26.3	21.1	29.3	43.9
Convergence Index (10)	3	6	9	48.8	44.1	16.5	19.8	32.3	24.4
Euclidian distance to tectonic lines	4	5	9	53.4	53.3	23.6	18.3	29.8	34.9
Euclidian distance to nappe boundaries	4	5	9	74.2	56.1	24.1	11.8	50.1	44.3
Convergence Index (50)	3	5	8	56.1	64.3	21.3	26.3	34.8	38.0
Void space (0–20 cm)	3	4	7	0.3	0.2	0.0	0.1	0.3	0.1
Catchment area (log)	2	4	6	43.4	45.0	4.9	10.2	38.6	34.8
Permeability (0–20 cm)	3	1	4	36.9	23.7	15.3	8.3	21.6	15.4
East aspect	2	1	3	9.4	6.5	5.2	4.8	4.2	1.8
North aspect	2	0	2	11.8	19.5	8.1	17.9	3.7	1.6
Void space (mean 0–100 cm)	0	2	2	0.0	0.0	0.0	0.0	0.0	0.0
Permeability (min)	0	0	0	1.1	4.9	0.5	0.4	0.6	4.5

NHESSD
1, 1001–1050, 2013

Assessing the quality of landslide susceptibility maps of Lower Austria

H. Petschko et al.

Table 3. Number of selected variables for the models GAM_1–GAM_{16}. Median AUROC values, transferability and thematic consistency of the modelling domains for spatial (spCV) and non-spatial (nspCV) cross-validation. Lower index values T_{sp}, T_{nsp} and C_{sp}, C_{nsp} indicate strong spatial and non-spatial transferability and thematic consistency, respectively.

Domain	Name	No. of selected variables	Median AUROC spCV	Median AUROC nspCV	Transferability index T_{sp}	Transferability index T_{nsp}	Consistency index C_{sp}	Consistency index C_{nsp}
239	Igneous rocks of the Austroalpine Unit	7	81.7	86.2	0.000	0.000	0.247	0.151
58	Molasse, Schlier	6	92.2	94.3	0.015	0.000	0.364	0.191
39	Molasse Zone	6	91.4	91.4	0.029	0.000	0.418	0.326
35	Debris, till	7	75	82.7	0.042	0.023	0.352	0.253
02297	Alluvial deposits including lakes and wetlands	7	97.9	99.3	0.042	0.000	0.339	0.389
179	Austroalpine Unit with limestone, marls and sandstone	8	79.5	81.2	0.046	0.000	0.176	0.109
126	Rheno-danubian Flysch Zone	11	84.5	84.1	0.051	0.002	0.099	0.180
104	Intramontane Basins	7	90.8	95.4	0.052	0.000	0.164	0.324
0124	Quaternary fluvial terrace and anthropogenic deposits	6	90.8	94.3	0.070	0.033	0.363	0.283
10	Loess, Loam	5	97.8	96.7	0.084	0.010	0.349	0.174
251259	Bohemian Massif including Fault Zone	6	83.4	92.5	0.091	0.000	0.377	0.251
120	Mélange Zone, Klippen Zone	8	68.1	80.7	0.098	0.000	0.210	0.081
191	Austroalpine Unit with dolostone	9	74.7	77	0.098	0.000	0.162	0.068
3786	Waschberg Zone including Bohemian Massif with sedimentary cover	4	88.4	88	0.170	0.102	0.424	0.235
276	Bohemian Massif, plutonic rock	4	85.6	93.4	0.208	0.073	0.419	0.323
230	Permo-Mesozoic rocks (overlying the Austroalpine igneous rocks)	6	52.4	79	0.397	0.045	0.519	0.319

NHESSD
1, 1001–1050, 2013

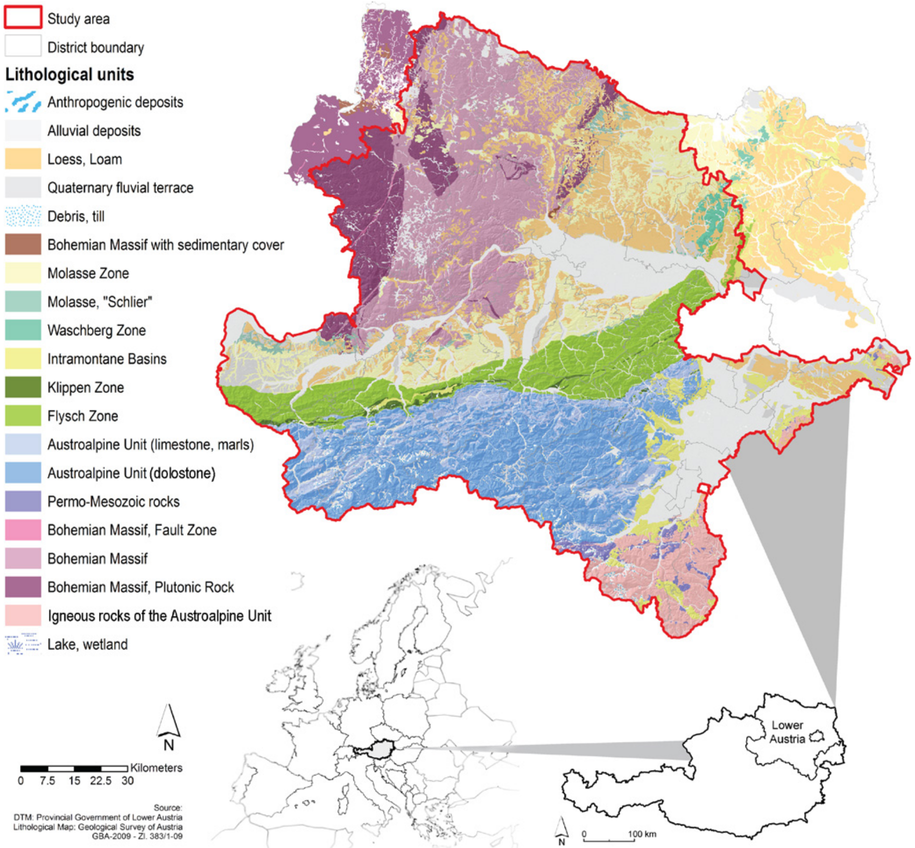

Fig. 1. Location and lithology of Lower Austria.

NHESSD

1, 1001–1050, 2013

Assessing the quality of landslide susceptibility maps of Lower Austria

H. Petschko et al.

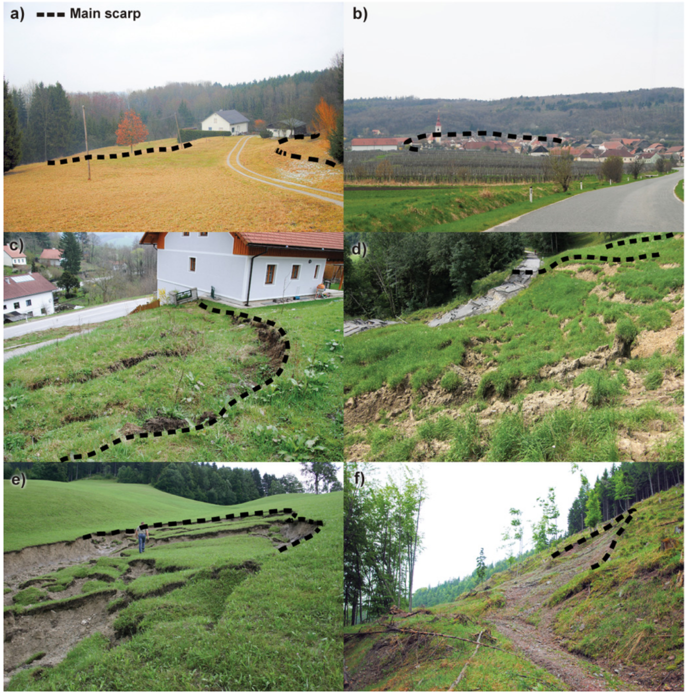

Fig. 2. Examples of landslides (earth and debris slides) typical of the respective lithological unit **(a)** landslide in Molasse Zone at Strengberg, **(b)** the village "Waitzendorf" located on a landslide in Molasse with Schlier, **(c)** translational landslide in the Rheno-danubian Flysch Zone at Brand **(d)** rotational landslide/flow in the Rheno-danubian Flysch Zone at Stössing, **(e)** landslide in Austroalpine Unit (Limestone, Marls) at Dippelreith, **(f)** landslide in Austroalpine Unit (with Dolostone) at Kleinzell. Pictures taken by: **(a)**, **(b)**, **(c)**, **(f)** Petschko (2012, 2010, 2011, 2012), **(d)** Bertsch (2009), **(e)** BGR (2006).

NHESSD
1, 1001–1050, 2013

Assessing the quality of landslide susceptibility maps of Lower Austria

H. Petschko et al.

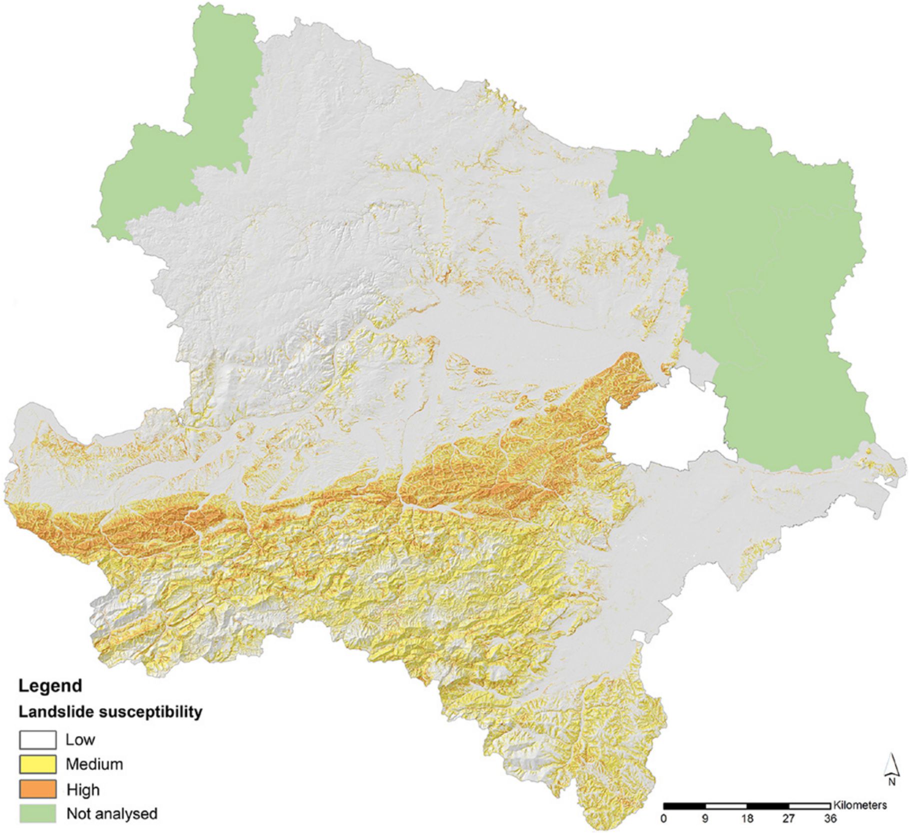

Fig. 3. Resulting classified landslide susceptibility map for Lower Austria. Data source: ALS hillshade – Provincial Government of Lower Austria.

NHESSD
1, 1001–1050, 2013

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◂ | ▸
◂ | ▸
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

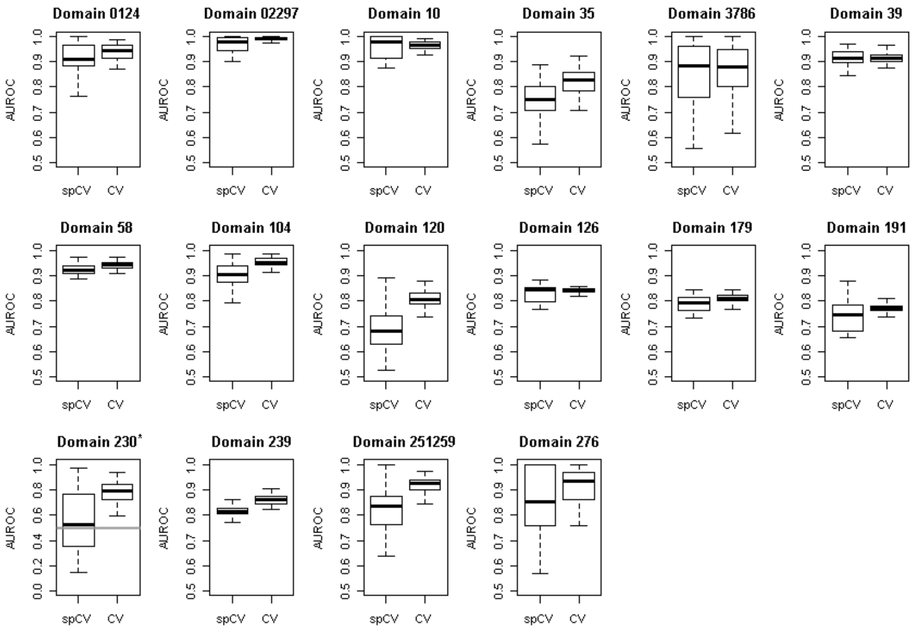

Fig. 4. Boxplots showing the range of AUROC values resulting from repeated *k*-fold cross-validation with spatial subsampling (spCV) and non-spatial subsampling (nspCV) for each modelling domain. * The y-axis limits for this domain range from 0 to 1 and a grey line marks the 0.5 value. In the other plots the y-axis limits range from 0.5 to 1 only.

NHESSD
1, 1001–1050, 2013

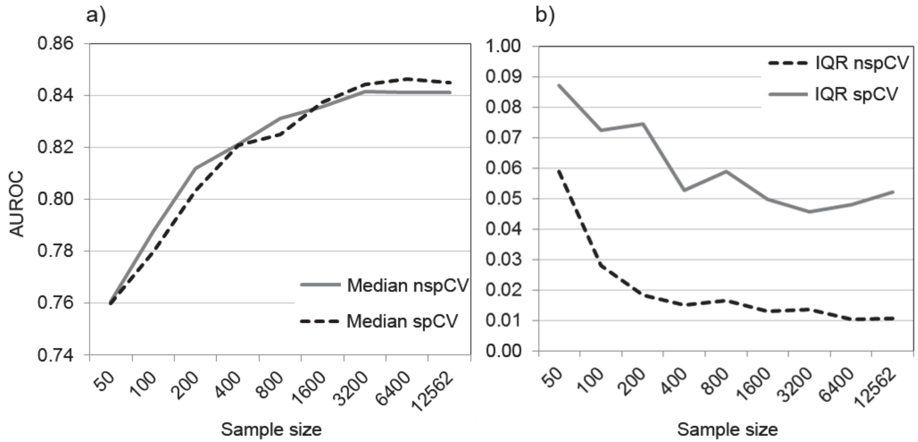

Fig. 5. **(a)** Median AUROC value and **(b)** interquartile range of AUROC values (IQR) at reduced sample sizes in the Flysch Zone (Domain 126) resulting from spatial cross-validation (spCV, grey line) and non-spatial cross-validation (nspCV, black dashed line).

NHESSD
1, 1001–1050, 2013

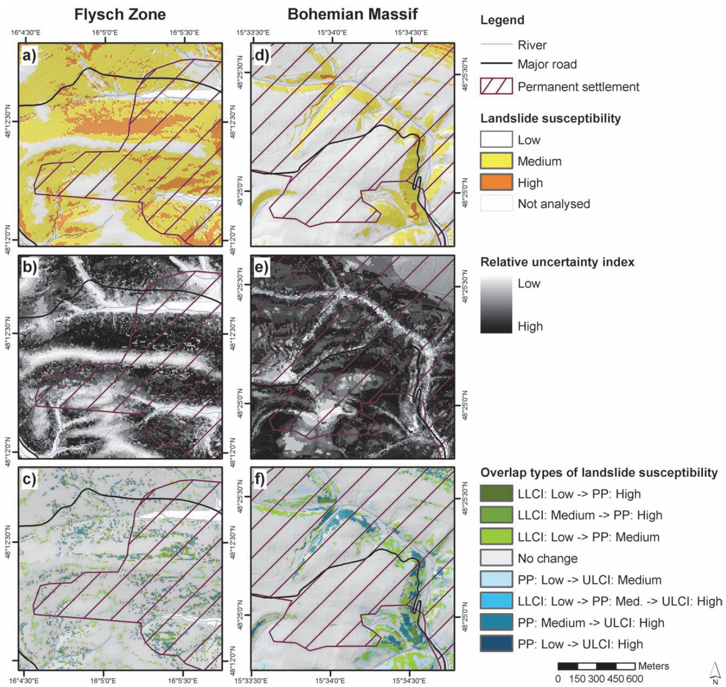

Fig. 6. Landslide susceptibility map (target scale 1 : 25 000) and map uncertainty in an example area. **(a, d)** susceptibility map; **(b, e)** relative uncertainty index; **(c, f)** susceptibility class uncertainty. **(a–c)** correspond to an area in the Flysch Zone (very high landslide density), **(d–f)** are in the Bohemian Massif (very low landslide density). The susceptibility class uncertainty refers to differences between susceptibility maps at the lower confidence limit (LLCI) or the upper confidence limit (ULCI) at the 95 % confidence level relative to the class in the predicted probability map (PP) in **(a)** and **(d)**. Data sources: ALS hillshade, river, major road and settlement – Provincial Government of Lower Austria.

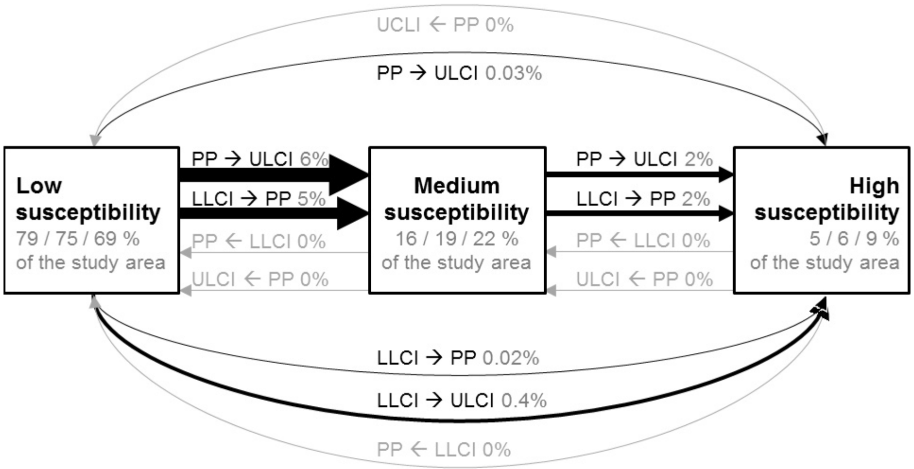

Fig. 7. Susceptibility class uncertainty expressed by types of overlaps between susceptibility maps at the lower confidence limit (LLCI) or the upper confidence limit (ULCI) at the 95 % confidence level relative to the class in the predicted probability map (PP). With the susceptibility class the percentage of area in the maps of LLCI, PP and ULCI is given in the box. The arrow thickness is relative to the percentage of affected area. Possible types of overlaps, which did not occur in the study area, are indicated with grey arrows. Next to the type of overlap the percentage of affected area related to the study area is given. In 85 % of the study area, the 95 % confidence limits fall within the same susceptibility class (not indicated here).

NHESSD
1, 1001–1050, 2013

Assessing the quality of landslide susceptibility maps of Lower Austria

H. Petschko et al.