Evaluating the Effect of Modelling Methods and Landslide Inventories Used for Statistical Susceptibility Modelling

27

Stefan Steger, Rainer Bell, Helene Petschko, and Thomas Glade

Abstract

Landslide susceptibility maps can be elaborated using a variety of methodological approaches. This study investigates quantitative and qualitative differences between two statistical modelling methods, taking into account the impact of two different response variables (landslide inventories) for the Rhenodanubian Flysch zone of Lower Austria. Quantitative validation of the four generated susceptibility maps is conducted by calculating conventional accuracy statistics for an independent random landslide subsample. Qualitative geomorphic plausibility is estimated by comparing the final susceptibility maps with hillshades of a high resolution Airborne Laser Scan Digital Terrain Model (ALS-DTM). Spatial variations between the final susceptibility maps are displayed by difference maps and their densities. Although statistical quality criterions reveal similar qualities for all maps, difference maps and geomorphic plausibility expose considerable differences between the maps. Given that, this conclusion could only be drawn by evaluating additionally the geomorphic plausibility and difference maps. Therefore, we indicate that conventional statistical quality assessment should be combined with qualitative validation of the maps.

Keywords

Landslide susceptibility • Geomorphic plausibility • Validation

27.1 Introduction

Landslides represent a widespread hazard for residents and their properties in many hilly and mountainous areas of the world. The relative spatial likelihood of a certain area possibly endangered by a landslide can be displayed by landslide susceptibility maps. In order to assess landslide susceptibility at medium and small scale, statistical susceptibility methods represent the most frequently used approaches (Cascini 2008; Van Westen 2000).

A literature review demonstrated that numerous studies focus on the statistical assessment of model performance computing confusion matrices, receiver-operating characteristic (ROC) plots, the area under the ROC-curve (AUROC) or prediction rates (Beguería 2006; Chung and Fabbri 2003; Frattini et al. 2010).

We assume that the explanatory power of these statistical measures is highly dependent on the quality of the inventory used to calculate these statistics.

However, previous studies outlined the need for qualitative evaluation methods such as the geomorphic plausibility (Bell 2007; Demoulin and Chung 2007). Therefore, we investigate the effect of different modelling methods and different inventories by validating the resulting maps statistically and qualitatively.

27.2 Study Area and Data

The Rhenodanubian Flysch zone of the provincial state Lower Austria is located in the eastern part of Austria and covers an area of 1,354 km^2. The west-to-east oriented

S. Steger (✉) · R. Bell · H. Petschko · T. Glade
Department of Geography and Regional Research, University of Vienna, Universitaetsstrasse 7, 1010, Wien, Austria
e-mail: stefan.steger@univie.ac.at

G. Lollino et al. (eds.), *Engineering Geology for Society and Territory – Volume 2*,
DOI: 10.1007/978-3-319-09057-3_27,

highly human influenced tectonic unit can be divided into five geologic sub-units (Fig. 27.1) and is well known to be susceptible to rainfall triggered landslides (Steger 2012).

The four susceptibility maps, which represent the basis for this study, were generated in an earlier work using variables (slope, geology, aspect, curvature, topographic position index, mean annual precipitation) selected on the basis of a comprehensive exploratory data analysis (Steger 2012). Susceptibility was calculated applying two methods: multivariate logistic regression (Atkinson and Massari 1998) and bivariate Landslide Susceptibility Index (LSI) (Lee 2004).

The two landslide inventories used to generate and validate the susceptibility maps quantitatively are the (i) highly accurate ALS-Inventory, which contains 6,218 landslide scarps and was mapped from a high resolution ALS-DTM purposely for susceptibility modelling (Petschko et al. 2013) and the (ii) less accurate building ground register (BGR), an archive consisting of 681 landslides representing the damaging events documented in this area. Furthermore, a hillshade was generated from a high-resolution ALS-DTM (1 m × 1 m) to assess the geomorphic plausibility.

27.3 Methods

Quantitative validation was conducted by comparing the generated susceptibility maps with 20 % of randomly selected landslides for the respective inventory. This independent subsample was not used to generate the maps. The output of the LSI-models (values ranging from 0 to ∞) did not allow to compute the "classical" threshold independent ROC-curves and AUROC-values. Therefore, four threshold dependent (logistic regression: 0.1, 0.3, 0.8, median; LSI: 1, 1.2, 1.5, median) confusion matrices were calculated for each model. The resulting sensitivity/specificity-pairs were plotted on a ROC-graph (Beguería 2006).

Based on the hillshade, geomorphic plausibility (Bell 2007) was subjectively assessed by evaluating the geomorphic situation of the area and by comparing this evaluation with equally classified (quartiles) susceptibility maps at medium scale. If this estimation converged with the ranking expressed in the susceptibility map, high geomorphic plausibility was assigned and vice versa.

Difference maps and their densities (values averaged over a radius of 500 m) were calculated to visualize pixel-based spatial variations between the equally classified susceptibility maps.

27.4 Results

The ROC-curves (Fig. 27.2e) reveal similar predictive capabilities for all susceptibility maps. A comparison of this curves with the associated AUROC-values of the logistic regression models (ALS: 0.83, BGR: 0.79) approves that all maps exhibit an acceptable to excellent ability to distinguish between susceptible and non-susceptible areas (Hosmer and Lemeshow 2000). In contrast, geomorphic plausibility differs substantially among these maps. Generally we observed that similar maps were created using the same inventory but different modelling methods. However, the maps differed widely using different inventories but the same modelling approach. Looking at the geomorphic situation in the representative section of the study area presented in Fig. 27.2 we expected the hummocky formed slopes of the northern transition zone (see Fig. 27.1) being a landslide prone area, whereas the smooth formed ridges and floodplains are not susceptible. Comparing this expert based evaluation with the susceptibility maps, we observed that ALS-maps resulting from both modelling methods (Fig. 27.2a, b) were frequently able to match our evaluation of the prevalent geomorphic situation. This tendency can be observed throughout the study area. Conversely, both BGR-maps (Fig. 27.2c, d) were regularly not able to

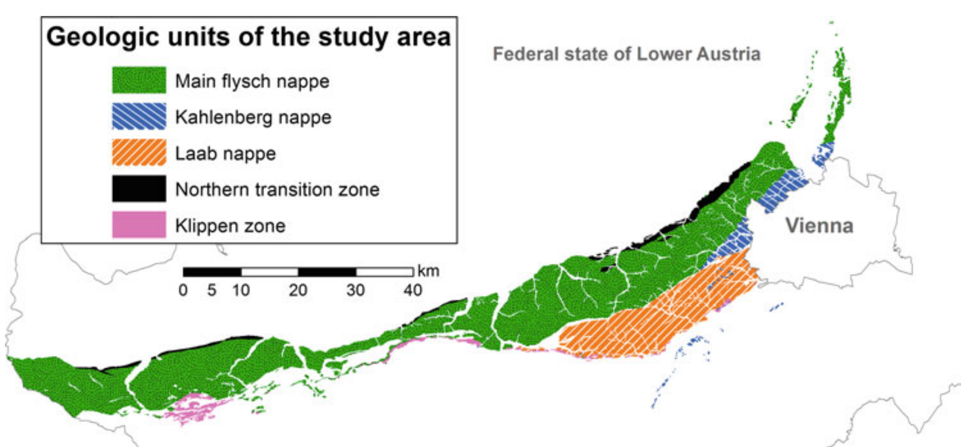


Fig. 27.1 Location and geologic overview of the study area

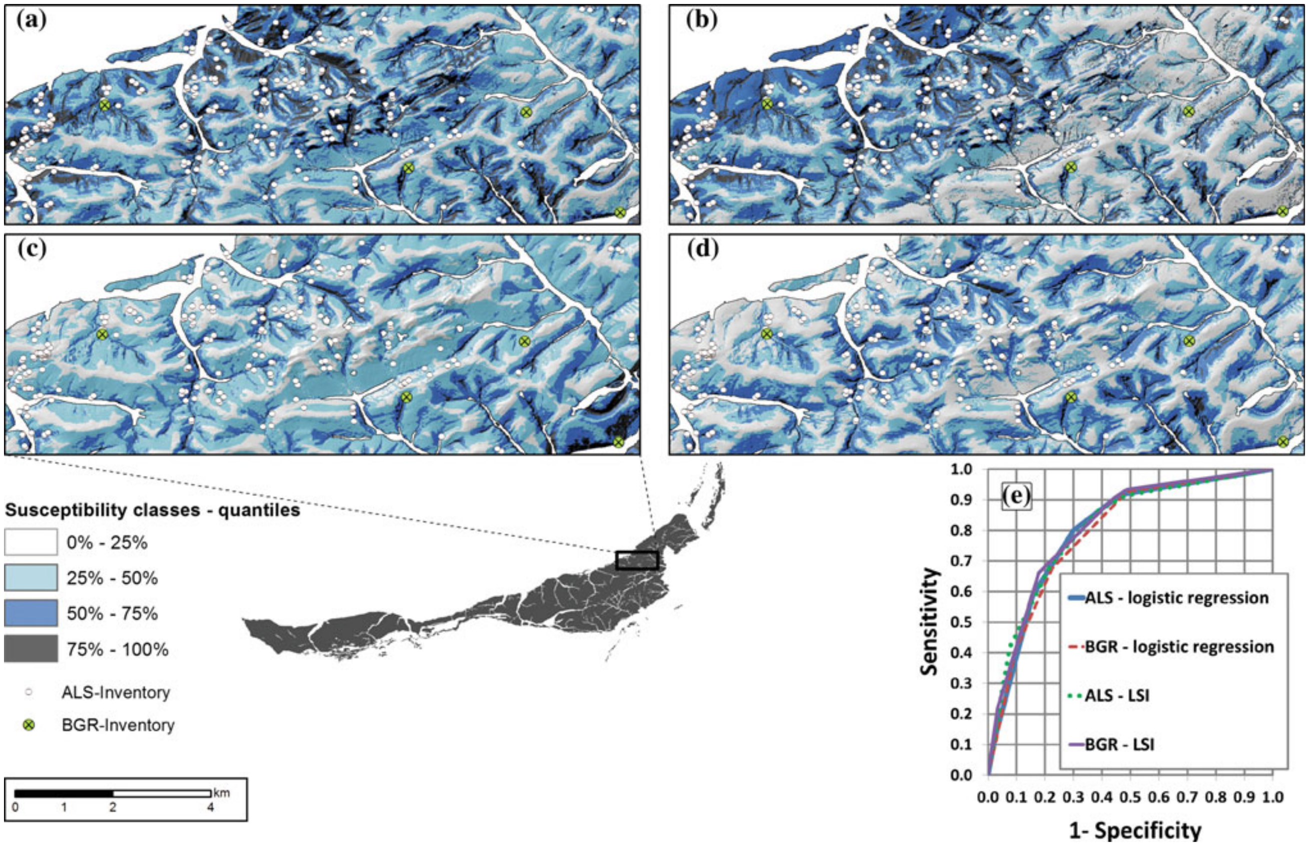


Fig. 27.2 ROC-curves (**e**) and susceptibility maps created by logistic regression (**a**, **c**) and LSI (**b**, **d**) using ALS-Inventory (**a**, **b**) and BGR-Inventory (**c**, **d**)

discriminate between susceptible and non-susceptible areas. In order to evaluate the effects of different modelling approaches on the geomorphic plausibility of the maps, we detected that both logistic regression maps (Fig. 27.2a, c) exhibit smoother transitions between the susceptibility classes as the more scattered LSI maps (Fig. 27.2b, d).

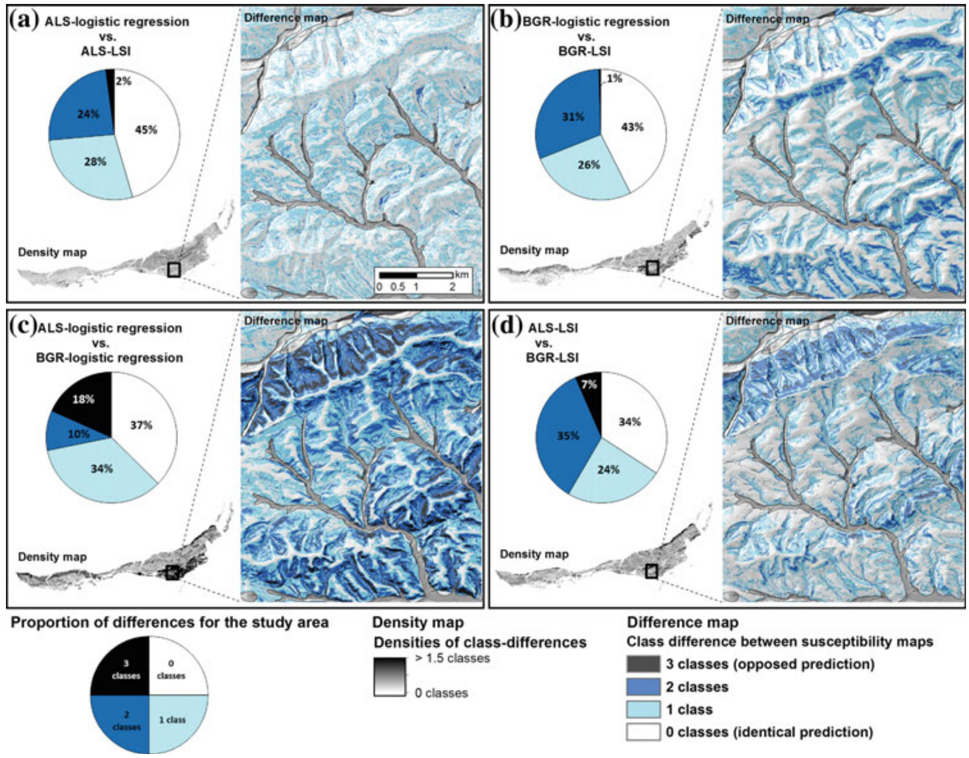
Difference maps and their densities (Fig. 27.3) highlight substantial spatial variations among the susceptibility maps. In this respect, largest dissimilarities were observed between maps created by different inventories (Fig. 27.3c, d). The density maps display the small-scale differences between the maps and point out that their main variations are frequently in spatial agreement with the areal extend of the geologic units (Fig. 27.1). Therefore, the differences can be explained by differing weightings of geologic units between the models. At medium scale (Fig. 27.3, difference maps) highest differences between the maps can be detected on steeper slopes and floodplains, whereas ridges are regularly displayed as non-susceptible by all maps.

The plots in Fig. 27.3 summarize the pixel-based proportion of differences between the maps for the entire study area. These charts reveal that even the most similar appearing maps (Fig. 27.3a) with similar statistical quality (Fig. 27.2e) and comparable geomorphic plausibility (Fig. 27.2a, b), differ largely by evaluating the differences at largest scale (pixel-basis). Accordingly, the proportion of dissimilar predictions between all maps is high and ranges from 55 (Fig. 27.3a) to 66 % (Fig. 27.3d). A relative high proportion (18 and 7 %) of completely opposed predictions was observed by comparing the maps created by the same modelling method using different inventories (Fig. 27.3c, d).

27.5 Discussion and Conclusion

The results of conventional statistical validation portrayed similar predictive capabilities for all generated susceptibility maps (Fig. 27.2e). Conversely, the assessment of geomorphic plausibility of these maps revealed substantial differences among the maps. Based on the results of this study and the knowledge of the differing qualities of both inventories (Petschko et al. 2013), we state that the explanatory power of inventory based statistical performance measures is only as high as the quality of the inventory used to calculate these statistics. Consequently, the discrepancy between the calculated high statistical quality (Fig. 27.2e) and the observed

Fig. 27.3 Density maps, difference maps and pixel-based proportion of differences displaying the spatial variation of calculated susceptibility maps at different scales



low geomorphic plausibility of both BGR-maps (Fig. 27.2c, d) can be explained. Since a complete highly accurate and unbiased landslide inventory is seldom available (Malamud et al. 2004; Petschko et al. 2013), we assume that a direct deduction of the reliability and applicability of these maps on the basis of conventional accuracy statistics may result in misleading conclusions.

Thus, we conclude that an (inventory-independent) assessment of geomorphic plausibility quality criterion should be performed additionally.

The implementation of susceptibility maps into spatial planning is accompanied by constructing a decisive reality for end-users (Petschko et al. 2014) which is often directly deduced from the respective color and/or pixel-value. Accordingly, the observed large differences at pixel-scale (Fig. 27.3) might result in major problems when implementing these maps. Further research on the geomorphic plausibility and the comparison of different modelling approaches is envisaged.

References

Atkinson PM, Massari R (1998) Comput Geosci 24:373

Beguería S (2006) Nat Hazards 37:315

Bell R (2007) Dissertation, Rheinische Friedrich-Wilhelms-Universität Bonn

Cascini L (2008) Eng Geol 102:164

Chung CJ, Fabbri AG (2003) Nat Hazards 30:451

Demoulin A, Chung C-JF (2007) Geomorphology 89:391

Frattini P, Crosta G, Carrara A (2010) Eng Geol 111:62

Hosmer DW, Lemeshow S (2000) Applied logistic regression. Wiley, New York

Lee S (2004) Environ Manag 34:223

Malamud BD, Turcotte DL, Guzzetti F, Reichenbach P (2004) Earth Surf Proc Land 29:687

Petschko H, Bell R, Leopold P, Heiss G, Glade T (2013) Landslide science and practice. Springer, Berlin, pp 281–286

Petschko H, Brenning A, Bell R, Goetz J, Glade T (2014) Nat Hazards Earth Syst Sci 14:95

Steger S (2012) Diploma thesis, University of Vienna

Van Westen CJ (2000) Surv Geophys 21:241