Geophysical Research Abstracts
Vol. 15, EGU2013-10276, 2013
EGU General Assembly 2013

Exploratory data analysis as an alternative to automated variable selection in statistical susceptibility modeling. A case study for the Flysch Zone (Lower Austria)

Stefan Steger, Rainer Bell, Helene Petschko, and Thomas Glade
Dept. of Geography and Regional Research, University of Vienna, Vienna, Austria (stefansteger@hotmail.com)

Statistical susceptibility modeling presents a widespread method to identify landslide prone areas. Since these quantitative methods are highly dependent on the quality of the input data, an appropriate selection of variables is essential. In this study a comprehensive spatial analysis of slope conditions, where landslides occurred in the rhenodanubian Flysch Zone of Lower Austria (a province in Austria), was performed to identify the most important factors which may lead to landslides in the study area. Analyses were conducted using two landslide inventories to demonstrate the impact of different dependent variables on the results. The ALS-Inventory was generated purposely for susceptibility modeling using a high resolution airborne laser scanning digital terrain model (ALS DTM) in a previous study. The building ground register (BGR) obtained by the Geological Survey of Lower Austria represents the damaging events of landslides in Lower Austria.
The results reveal the relationships between the spatial occurrence of landslides and its preparatory factors for the respective inventory. Based on this information explanatory variables were selected. Only variables providing an explanation for the spatial distribution of landslides (for the respective spatial resolution of input data) were considered. Thus the limitations of input data were uncovered and taken into account in the statistical modeling. Although the study area is highly affected by human impact, the interpretation of the results lead to the exclusion of all land use - variables from subsequent modeling. In fact it was discovered that an inclusion of the land cover variables leads to unrealistic susceptibility values. The explanatory data analyses lead to the selection of six variables for the ALS-inventory and five variables for the BGR out of 19 independent variables. For the susceptibility modeling two statistical approaches, namely multivariate logistic regression analysis and a bivariate index-based approach (Landslide Susceptibility Index, LSI) were applied on both inventories in separate calculations. Quantitative quality assessment results revealed similar levels of statistical quality for all models. However, differing geomorphological plausibility of obtained susceptibility maps indicate that high statistical quality alone does not guarantee high qualitative modeling results. Furthermore it was shown, that in this case the modeling method used has a lower impact on the quality of the results than the landslide inventories used. The results clearly emphasize the importance of an appropriate input data selection which can be assessed by exploratory data analyses. Therefore a comprehensive preparatory data analysis can be considered as a serious alternative to purely automated variable selection for statistical susceptibility models.