Geophysical Research Abstracts
Vol. 17, EGU2015-3220, 2015
EGU General Assembly 2015

Selecting statistical or machine learning techniques for regional landslide susceptibility modelling by evaluating spatial prediction

Jason Goetz (1), Alexander Brenning (1), Helene Petschko (1), and Philip Leopold (2)

(1) Department of Geography, Friedrich Schiller University of Jena, Jena, Germany, (2) Health and Environment Department, Austrian Institute of Technology GmbH, Tulln, Austria

With so many techniques now available for landslide susceptibility modelling, it can be challenging to decide on which technique to apply. Generally speaking, the criteria for model selection should be tied closely to end users' purpose, which could be spatial prediction, spatial analysis or both. In our research, we focus on comparing the spatial predictive abilities of landslide susceptibility models. We illustrate how spatial cross-validation, a statistical approach for assessing spatial prediction performance, can be applied with the area under the receiver operating characteristic curve (AUROC) as a prediction measure for model comparison. Several machine learning and statistical techniques are evaluated for prediction in Lower Austria: support vector machine, random forest, bundling with penalized linear discriminant analysis, logistic regression, weights of evidence, and the generalized additive model. In addition to predictive performance, the importance of predictor variables in each model was estimated using spatial cross-validation by calculating the change in AUROC performance when variables are randomly permuted. The susceptibility modelling techniques were tested in three areas of interest in Lower Austria, which have unique geologic conditions associated with landslide occurrence.

Overall, we found for the majority of comparisons that there were little practical or even statistically significant differences in AUROCs. That is the models' prediction performances were very similar. Therefore, in addition to prediction, the ability to interpret models for spatial analysis and the qualitative qualities of the prediction surface (map) are considered and discussed. The measure of variable importance provided some insight into the model behaviour for prediction, in particular for "black-box" models. However, there were no clear patterns in all areas of interest to why certain variables were given more importance over others.