EGU2020-8365, updated on 02 Aug 2021
https://doi.org/10.5194/egusphere-egu2020-8365
EGU General Assembly 2020

Random forest classification of morphology in the northern Gerecse (Hungary) to predict landslide-prone slopes

Gáspár Albert[1] and Dávid Gerzsenyi[2]
[1]Eötvös Loránd University, Cartography and Geoinformatics, Budapest, Hungary (albert@ludens.elte.hu)
[2]Eötvös Loránd University, Doctorate School of Earth Sciences, Budapest, Hungary (gerzsd@caesar.elte.hu)

The morphology of the Gerecse Hills bears the imprints of fluvial terraces of the Danube River, Neogene tectonism and Quaternary erosion. The solid bedrocks are composed of Mesozoic and Paleogene limestones, marls, and sandstones, and are covered by 115 m thick layers of unconsolidated Quaternary fluvial, lacustrine, and aeolian sediments. Hillslopes, stream valleys, and loessy riverside bluffs are prone to landslides, which caused serious damages in inhabited and agricultural areas in the past. Attempts to map these landslides were made and the observations were documented in the National Landslide Cadastre (NLC) inventory since the 1970's. These documentations are sporadic, concentrating only on certain locations, and they often refer inaccurately to the state and extent of the landslides. The aim of the present study was to complete and correct the landslide inventory by using quantitative modelling. On the 480 sq**.** km large study area all records of the inventory were revisited and corrected. Using objective criteria, the renewed records and additional sample locations were sorted into one of the following morphological categories: scarps, debris, transitional area, stable accumulation areas, stable hilltops, and stable slopes. The categorized map of these observations served as training data for the random forest classification (RFC).

Random forest is a powerful tool for multivariate classification that uses several decision trees. In our case, the predictions were done for each pixels of medium-resolution (~10 m) rasters. The predictor variables of the decision trees were morphometric and geological indices. The terrain indices were derived from the MERIT DEM with SAGA routines and the categorized geological data is from a medium-scale geological map [1]. The predictor variables were packed in a multi-band raster and the RFC method was executed using R 3.5 with RStudio.

After testing several combinations of the predictor variables and two different categorisation of the geological data, the best prediction has cca. 80% accuracy. The validation of the model is based on the calculation of the rate of well-predicted pixels compared to the total cell-count of the training data. The results showed that the probable location of landslide-prone slopes is not restricted to the areas recorded in the National Landslide Cadastre inventory. Based on the model, only ~6% of the estimated location of the highly unstable slopes (scarps) fall within the NLC polygons in the study area.

The project was partly supported by the Thematic Excellence Program, Industry and Digitization

Subprogram, NRDI Office, project no. ED_18-1-2019-0030 (from the part of G. Albert) and the ÚNKP-19-3 New National Excellence Program of the Ministry for Innovation and Technology (from the part of D. Gerzsenyi).

Reference:

[1] Gyalog L., and Síkhegyi F., eds. Geological map of Hungary (scale: 1:100 000). Budapest, Hungary, Geological Institute of Hungary, 2005.