Sensitivity analysis and scale issues in landslide susceptibility mapping

Geophysical Research Abstracts
Vol. 15, EGU2013-7855, 2013
EGU General Assembly 2013

Sensitivity analysis and scale issues in landslide susceptibility mapping

Filippo Catani, Daniela Lagomarsino, Samuele Segoni, and Veronica Tofani

University of Firenze - Dipartimento di Scienze della Terra, Earth Sciences Department, Firenze, Italy (filippo.catani@unifi.it, +39 055 2757599)

Despite the large number of recent advances and developments in landslide susceptibility mapping (LSM) there is still a lack of studies focusing on specific aspects of LSM model sensitivity. For example, the influence of factors of paramount importance such as the survey scale of the landslide conditioning variables (LCVs), the resolution of the mapping unit (MUR) and the optimal number and ranking of LCVs have never been investigated analytically, especially on large datasets.

In this paper we attempt this experimentation concentrating on the impact of model tuning choice on the final result, rather than on the comparison of methodologies. To this end, we adopt a simple implementation of the random forest (RF) classification family to produce an ensamble of landslide susceptibility maps for a set of different model settings, input data types and scales. RF classification and regression methods offer a very flexible environment for testing model parameters and mapping hypotheses, allowing for a direct quantification of variable importance. The model choice is, in itself, quite innovative since it is the first time that such technique, widely used in remote sensing for image classification, is used in this form for the production of a LSM.

Random forest is a combination of tree (usually binary) bayesian predictors that permits to relate a set of contributing factors with the actual landslides occurrence. Being it a nonparametric model, it is possible to incorporate a range of numeric or categorical data layers and there is no need to select unimodal training data. Many classical and widely acknowledged landslide predisposing factors have been taken into account as mainly related to: the lithology, the land use, the land surface geometry (derived from DTM), the structural and anthropogenic constrains. In addition, for each factor we also included in the parameter set the standard deviation (for numerical variables) or the variety (for categorical ones).

The use of random forest enables to estimate the relative importance of the single input parameters and to select the optimal configuration of the regression model. The model was initially applied using the complete set of input parameters, then with progressively smaller subsamples of the parameter space. Considering the best set of parameters we also studied the impact of scale and accuracy of input variables and the influence of the RF model random component on the susceptibility results.

We apply the model statistics to a test area in central Italy, the hydrographic basin of the Arno river (ca. 9000 km2), we present the obtained results and discuss them. We also use the outcomes of the parameter sensitivity analysis to investigate the different role of environmental factors in the test area.