How complete should be the landslide inventory to generate a reliable landslide susceptibility model?

Sérgio C. Oliveira, José L. Zêzere, Sandra Oliveira, and Ricardo A.C. Garcia

Universidade de Lisboa, Centro de Estudos Geográficos, IGOT, Lisboa, Portugal

Historical landslide inventory maps are frequently incomplete and this is usually pointed out as a source of uncertainty affecting the predictive ability of data-driven landslide susceptibility models. Nevertheless, the concept of having a complete landslide inventory seems far from reaching a consensual definition, both from the theoretical and practical point of view. Landslide mapping over time depends on a wide range of factors, some of them as simple as the ambitious capability to regionally map all the landslide features/signatures, detectable by fieldwork or image interpretation, which are often lost in a short period due to erosion or man intervention, which erase the morphological signs of instability.

Within this framework, the main goal of the present work is to assess to what extent a presumed complete shallow slides inventory map is necessary to consistently assess the susceptibility to shallow slides in a certain area. The working hypothesis is tested in the Grande da Pipa river (GPR) basin, which extends for 110 square kilometers in the north of Lisbon region, Portugal. To assess susceptibility to shallow slides occurrence, we apply a bivariate statistical method (the Information Value), using an inventory containing more than 500 shallow slides ranging in size from 10 to 8000 square meters; and a dataset of eight terrain predisposing factors (lithology, slope, aspect, plan curvature, slope area ratio, topographic position index, soil type and land use), supported by automated R routines. In a first moment, the modelling strategy encompasses the creation of three independent blocks of landslide cases to be used for training (70 % of the landslides) and validation (30 % of the landslides) based on a random partition of the shallow slides inventory. In a second phase, each training group, for the different blocks, is randomly split in 14 new landslide sample groups, through a progressive increment of 5 % in the number of landslide cases included, from sample group 1 (with only 5 % of the total landslides) to sample group 14 (with 70 % of the total landslides), to obtain the shallow slides susceptibility scores. The validation of each of the 14 susceptibility map, from the different blocks, is done independently with the validation group of shallow slides (30 %) previously set aside and not used for susceptibility modelling. To accomplish that, we graphically compute prediction rate curves and calculate the respective area under the curve.