Geomorphology 253 (2016) 508–523

Contents lists available at ScienceDirect

Geomorphology

journal homepage: www.elsevier.com/locate/geomorph

Different landslide sampling strategies in a grid-based bi-variate statistical susceptibility model

Haydar Y. Hussin [a,*], Veronica Zumpano [b], Paola Reichenbach [c], Simone Sterlacchini [d], Mihai Micu [b], Cees van Westen [a], Dan Bălteanu [b]

[a] Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, 7500 AE, Enschede, The Netherlands
[b] Institute of Geography, Romanian Academy, 12 Dimitrie Racovita, 023993 Bucharest, Romania
[c] National Research Council, Research Institute for Geo-Hydrological Protection, Via della Madonna Alta 126, 06128 Perugia, Italy
[d] National Research Council, Institute for the Dynamics of Environmental Processes, Piazza della Scienza 1, 20126 Milano, Italy

ARTICLE INFO

Article history:
Received 18 May 2015
Received in revised form 28 October 2015
Accepted 29 October 2015
Available online 9 November 2015

Keywords:
Landslide susceptibility
Weights-of-evidence (WofE)
Landslide sampling
Grid-based analysis

ABSTRACT

This study had three aims. The first was to assess the performance of the weights-of-evidence (WofE) landslide susceptibility model in areas that are very different in terms of size, geoenvironmental settings, and landslide types. The second was to test the appropriate strategies to sample the mapped landslide polygon. The final aim was to evaluate the performance of the method to changes in the landslide sample size used to train the model. The method was applied to two areas: the Fella River basin (eastern Italian Alps) containing debris flows, and Buzau County (Romanian Carpathians) with shallow landslides. The three landslide sampling strategies used were: (1) the landslide scarp centroid, (2) points populating the scarp on a 50-m grid, and (3) the entire scarp polygon. The highest success rates were obtained when sampling shallow landslides as 50-m grid-points and debris flow scarps as polygons. Prediction rates were highest when using the entire scarp polygon method for both landslide types. The sample size test using the landslide centroids showed that a sample of 104 debris flow scarps was sufficient to predict the remaining 941 debris flows in the Fella River basin, while 161 shallow landslides was the minimum required number to predict the remaining 1451 scarps in Buzau County. Below these landslide sample thresholds, model performance was too low. However, using more landslides than the threshold produced a *plateau effect* with little to no increase in the model performance rates.

1. Introduction

The spatial prediction of landslides in the form of susceptibility assessment studies have been applied now for the past 30 years with new techniques continuously being developed and updated. An overwhelming amount of literature has been published on the different methods that have been used throughout the years. The extensive guidelines, reviews, and overviews related to landslide hazard and risk (Varnes, 1984; Soeters and van Westen, 1996; van Westen et al., 1997, 2008; Aleotti and Chowdhury, 1999; Guzzetti et al., 1999; van Westen, 2000; Dai et al., 2002; Crozier and Glade, 2005; Glade and Crozier, 2005; Wang et al., 2005; Fell et al., 2008; Corominas et al., 2013) generally divide landslide susceptibility methods into qualitative (e.g., heuristic, geomorphological analysis, expert-based index/weighting) or (semi-) quantitative approaches (e.g., statistical and deterministic analysis). The quantitative statistical methods follow a single important assumption, that slope instability factors causing landslides in the past will statistically determine the spatial probability of landslide occurrence in the future (Soeters and van Westen, 1996). According to this assumption, the predictive capability of statistical susceptibility methods relies on two input data: the inventory of past landslide events and the landslide causative factor maps (also called *landslide predisposing* factors, *landslide conditional* factors, or *slope instability* factors). The way in which landslides are represented and sampled in a GIS determines how the causative factor information is extracted for susceptibility mapping and is therefore a very important aspect in landslide hazard zonation studies.

Landslides are generally mapped using vector-based representations of the landslide data, which are represented by points (Brenning, 2005; Galli et al., 2008), polygons (van Westen et al., 2000; Chung and Fabbri, 2005), and lines (Donati and Turrini, 2002). In some cases, slope failures can be directly mapped as raster data, for example by semiautomated mapping from remote-sensing imagery (Brenning, 2009; Mondini et al., 2011). The mapping representation is determined by the type and availability of data, the spatial scale of the analysis, the purpose of the study, and the mapping methods used, among others (Soeters and van Westen, 1996; Guzzetti et al., 1999; van Westen, 2004; Glade and Crozier, 2005; van Westen et al., 2008). All statistical landslide susceptibility zonations require the selection of mapping units, which are the

---

* Corresponding author.
E-mail address: h.y.hussin@utwente.nl (H.Y. Hussin).

http://dx.doi.org/10.1016/j.geomorph.2015.10.030

subdivisions that make up the susceptibility map. A variety of mapping units are reported in the literature (Guzzetti et al., 1999; Van Den Eeckhaut et al., 2009). The choice of the mapping unit is crucial because it also determines how landslides will be sampled to prepare the training and prediction (validation) subsets for the susceptibility modeling that can be vector-based (Carrara et al., 1995; Guzzetti et al., 2005; Galli et al., 2008) or grid-based (Carrara, 1983; van Westen, 1993; Chung and Fabbri, 1999; Remondo et al., 2003).

In grid-based (also referred to as pixel or raster-based) susceptibility assessments, landslide mapping representations are either overlaid in their original format (e.g., points, polygons) on grid-cell causative factor maps to directly extract data from the factor maps or are converted to a raster map and then used for data extraction. According to the literature concerning grid-based landslide susceptibility mapping, four general strategies are used to sample landslide pixels:

(1) The landslide is sampled as a single pixel (Atkinson and Massari, 1998; Van Den Eeckhaut et al., 2006; Thiery et al., 2007; Yilmaz, 2010; Piacentini et al., 2012). Usually, the pixel is the centroid of the entire landslide or the scarp area. The single pixel can be selected to represent the *top-point* of a landslide placed by an expert on the initiation area, which is not necessarily the centroid (Qi et al., 2010; Gorum et al., 2011; Xu et al., 2014). The single pixel is often applied if landslides have been mapped directly as points or if the landslides in polygon format are not reliable for the susceptibility analysis (e.g., data scarcity, size of the area, scale-related issues, etc.). When selecting a single grid-cell to represent a landslide, the rest of the mapping units (grid-cells) that could still be located within a landslide polygon are considered nonlandslide areas.

(2) All the pixels within the entire landslide body or the scarp area can be sampled as landslide pixels (Ayalew and Yamagishi, 2005; Poli and Sterlacchini, 2007; Blahut et al., 2010; Sterlacchini et al., 2011; Regmi et al., 2013; Petschko et al., 2014). In this case, all pixels located outside the landslide polygons are considered as nonlandslide areas.

(3) The main scarp upper edge (MSUE) approach selects pixels on and around the landslide crown-line (Donati and Turrini, 2002; Clerici et al., 2006;Jurko et al., 2006), which basically is the upper edge of the landslide scarp area. The MSUE method was applied for the following reasons (Donati and Turrini, 2002; Clerici et al., 2006): the upper edge of the scarp area was the most identifiable feature in the landslide mapping, the entire depletion zone (scarp area) was less visible owing to recovery of the slope, and the scarp area was often partly covered by the accumulation zone, making the boundary between the two zones difficult to identify. Similar to the seed-cell methodology, the MSUE method is able to represent the landslide using pixels in *undisturbed morphological conditions* by projecting an artificial crown-line at a certain distance from the original crown-line, with the distance and length assigned by the expert (Clerici et al., 2006).

(4) The *seed-cell* approach (Süzen and Doyuran, 2004; Nefeslioglu et al., 2008; Yilmaz, 2010; Demir et al., 2013; San, 2014) selects pixels within a buffer polygon around the upper landslide scarp area and sometimes part of the flanks of the accumulation zone. The buffer distance, which determines the number of cells representing the landslide, is defined by an expert. The purpose of this method according to Süzen and Doyuran (2004) is to consider 'that the best undisturbed morphological conditions (conditions before landslide occurrence) would be extracted from the vicinity of the landslide polygon itself'. However, this could lead to problems in cases where landslide boundaries coincide with main morphological boundaries (e.g., top of the landslide at the crest of a ridge).

A number of studies have compared the effect of different sampling strategies applied to landslide susceptibility zonation. Poli and Sterlacchini (2007) studied the landslide centroid and a selection of points populating the polygon every 50 and 20 m. They found that one point every 50 m within a landslide polygon performed better than using the single centroid or the 20-m points. Yilmaz (2010) compared the susceptibility using the scarp polygon, seed cells and a single point. According to Yilmaz (2010), 'validations of the obtained maps indicated that the more realistic results obtained from the analyses were obtained using the scarp sampling strategy, however, it was relatively similar with the seed cells strategy. It can be evaluated that the two strategies, such as scarp and seed cells considered, have relatively similar accuracy'. The single point sampling had lower performance rates. Simon et al. (2013) compared the extraction of slope angle information between landslide polygons and their centroids. They concluded that using centroid points could have some disadvantages, such as abstracting landslide causative information not located at the actual initiation points but located in less significant factor classes or even outside the actual polygon boundary because of using the point of gravity.

Once the expert determines which grid-cells are considered landslides or nonlandslide areas, a selection procedure is required to define the sampling size of pixels that will be exploited to train and validate the susceptibility model. The modeler needs to decide not only the number of landslide pixels but also the number of nonlandslide pixels to be used in assessing the success and prediction capability of the model. The ratio between landslide and nonlandslide areas depends, among others, on the type of statistical model used in the susceptibility assessment. As Heckmann et al. (2014) summarized for logistic regression and other types of regression analysis, the ratio often ranges between 1:1 and 1:10. However, larger ratios have also been used (Melchiorre et al., 2008; Felicísimo et al., 2013; Heckmann et al., 2014), including in other types of statistical techniques, such as the Bayesian approaches, where sometimes all the nonlandslide pixels are applied in the analysis (Blahut et al., 2010; Regmi et al., 2010).

Recent studies have been conducted to understand the effects of landslide sample size on susceptibility mapping and prediction (Hjort and Marmion, 2008; Heckmann et al., 2014; Petschko et al., 2014). Hjort and Marmion (2008) assessed the effect of sample size on the susceptibility of geomorphological processes such as permafrost and solifluction in an area of 600 km^2 using model resolutions of 1 and 25 ha. They found that for a sufficient model performance, producing AUC values ranging between 0.80 and 0.95, 100 to 200 samples were required of a population of more than 1700 data points. Heckmann et al. (2014) sampled 1000 nonlandslide subsets ranging the sample size from 50 to 5000 pixels of 5-m resolution in two small areas of 7 and 19 km^2, while sampling 81 landslide pixels. They recommended a minimum of 300–350 nonlandslide pixels, corresponding to a ratio of 1:3.7 to 1:4.3 (81:300–81:350) and obtaining an average area under the ROC curve of 0.83. Petschko et al. (2014) applied a 1:1 ratio of landslide to nonlandslide pixels of 5-m resolution in an area of 15,850 km^2 and found that as the sample size increased from 50 to 12,562 pixels (total number of landslides), so did the AUC of the ROC curve from 0.76 to 0.84, with a slight plateauing at 3200 pixels or 25% of the landslide inventory. The literature indicates that no ideal fixed percentage or ratio exists between landslide and nonlandslide sample sizes and is further dependent on the statistical technique used in the susceptibility analysis.

Most of the research analyzing the effects of landslide sampling strategies and landslide sample sizes on susceptibility mapping have either used regression analysis techniques (e.g., logistic, linear, multivariate regression, etc.) or machine learning methods (e.g., artificial neural networks, generalized boosting method, etc.). Furthermore, these works were conducted in single case study areas and mainly using single landslide types. Despite the WofE method being widely applied, the influence of landslide sample sizes in training the model, and the subsequent effect on performance and prediction rates, has rarely been conducted.