

Weighting predisposing factors for shallow slides susceptibility assessment at the regional scale

J.L. Zêzere, S.C. Oliveira, R.A.C. Garcia & E. Reis
Geographical Research Centre, University of Lisbon, Portugal

ABSTRACT: The quality of landslide prediction models at the regional scale does not automatically increase with the increasing number of layers assumed to be landslide predisposing factors. Additionally, the significance of such layers as predisposing factors is frequently not evaluated. In this paper we apply a sensitivity analysis to a statistically-based landslide susceptibility model, performed for shallow translational slides occurred in a test site located north of Lisbon (Portugal). The model is applied individually to each layer (e.g. slope, aspect, transverse slope profile, geomorphology, lithology, superficial deposits and land use) and to different combinations of overlapped layers. The computation of success-rates and prediction-rates for such models allows concluding that the relationship between number of variables within the prediction model and the quality of predicted results is not linear, and it is possible to obtain an accurate landslide susceptibility map using a limited number of instability predisposing factors in the prediction model.

1 INTRODUCTION

The assessment of susceptibility associated to mass movements reveal in recent years significant improvements in indirect statistically-based methods (Guzzetti et al. 1999, 2005, Zêzere et al. 2004). Current Spatial Data Analysis (SDA) techniques allow the independent validation of results in post-processing operations, for prediction models based on both bivariate and multivariate statistical methods (Fabbri et al. 2002, Chung & Fabbri 2005). Therefore, validation is not anymore exclusively dependent on the occurrence of new instability events.

Assessment of landslide susceptibility is always based on the assumption that future mass movements are more probable to occur in areas with conditions similar to those that originate slope instability in the past (Carrara et al. 1999). In this context, recent developments in Geographical Information Systems allow the development of models resulting from the spatial relationships between landslides and an increasing number of landslide predisposing factors. The quality of the landslide inventory and of the landslide predisposing factors database are of crucial importance for the quality of prediction results, independently on the statistical tools used for the modelling procedure. Usually, it is not easy to obtain systematic and detailed cartographic data that reflects directly the physical parameters involved in

slope instability (e.g., shearing forces, soil shear strength, and spatial and temporal variation of pore water pressure). Therefore, it is common to make recourse to the available cartography that may correlate with landslide distribution (e.g., terrain morphology, geology, land use). However, the significance of such themes as landslide predisposing factors is frequently not evaluated.

The present study aims at evaluating quantitatively the relevance of different predisposing layers for shallow translational slides susceptibility assessment, using a spatial data set from a test site located north of Lisbon (Portugal).

The main objectives of this study are: (i) to evaluate the relation between the number of variables within a statistic/probabilistic landslide susceptibility model and the quality of predicted results; and (ii) to assess the weight of each individual landslide predisposing factor by applying a sensitivity analysis, and to define the best variable combination by computing the corresponding success and prediction rates.

2 STUDY AREA

The sensitivity evaluation of landslide susceptibility models to the type and number of landslide predisposing factors was performed in a test site of 20 km² located northward of Lisbon. The test site of Fanhões – Trancão (Fig. 1) is characterized by the

monocline geological structure, and the layers dip from 5° to 25° towards south and southeast. From the lithological point of view, the outcropping rocks are very heterogeneous and include conglomerate, sandstone, claystone, marl, marly limestone, limestone, compacted basalt and volcanic tuff, dated from the Cretaceous to the Palaeogene.

The monocline setting and the diversity of geological formations sustain a *cuesta* landscape, and the Fanhões-Trancão test site is located in the dip slope of the *cuesta*, i.e., a substructural slope defined by a general coincidence between the topographical surface and the dip of the strata. The geological setting also controls the fluvial system, and the most important rivers run in the same direction of the dip of strata. This is the case of the Fanhões river and of the Trancão river, located in the west and east side of the test site, respectively (Fig. 1). The Fanhões and Trancão valleys are the most relevant geomorphologic features within the study area because of their strong deep and the general steep slopes, although the altitude of the area does not exceed 335 m.

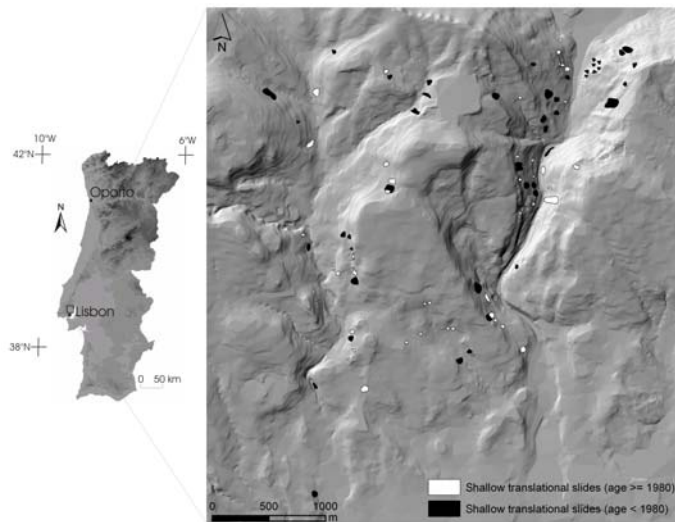


Figure 1. Study area location and distribution of shallow translational slides.

The detailed geomorphologic mapping (scale 1:2000) of the study area allowed the identification and characterization of 100 shallow translational slides, resulting in 143,000 m² of unstable area and corresponding to 0.7 % of the total study area. Shallow translational slides within the test site affect almost exclusively colluvium deposits and have minor dimension (mean area, 1422 m²; mean volume, 357 m³). These landslides occur along planar rupture zones located usually from 0.5 to 1.5 m below the topographic surface.

Shallow translational slides occurred in the study area during the last 4 decades have been triggered by intense rainfall periods ranging from 1 to 15 days (Zêzere & Rodrigues 2002, Zêzere et al. 2005). Intense rainfall is responsible by the rapid growth of pore pressure and by the loss of the apparent cohe-

sion of thin soils, resulting in failure within the soil material or at the contact with the underlying impermeable bedrock composed of claystone, marl or volcanic tuff (Zêzere et al. 2005, 2007).

3 LANDSLIDE SUSCEPTIBILITY ASSESSMENT AND SENSITIVITY ANALYSIS

The susceptibility assessment to shallow translational slides occurrence is based on the favorability concept (Chung & Fabbri 1993, Fabbri et al. 2002). Within this concept, we assume that future probability of landslide occurrence can be quantitatively evaluated by bi-variate statistical relationships between the spatial distribution of past landslides and several types of independent spatial data sets that are understood as landslide predisposing factors.

The landslide predisposing factors used in this study are the following: slope angle, slope aspect, transversal slope profile, lithology, superficial deposits, geomorphological units and land use (Table 1). More details about the data collection and database structure can be found in Reis et al. (2003) and Zêzere et al. (2004, 2007).

Table 1. Predisposing factors used for shallow translational slides susceptibility assessment.

Id	Predisposing factor	# of classes
A	Slope angle	8
B	Slope aspect	9
C	Transversal slope profile	5
D	Lithological units	6
E	Superficial deposits	7
F	Geomorphological units	11
G	Land use	6

Figure 2 summarizes the methodological procedures for the landslide susceptibility assessment and validation, as well as for the sensitivity analysis.

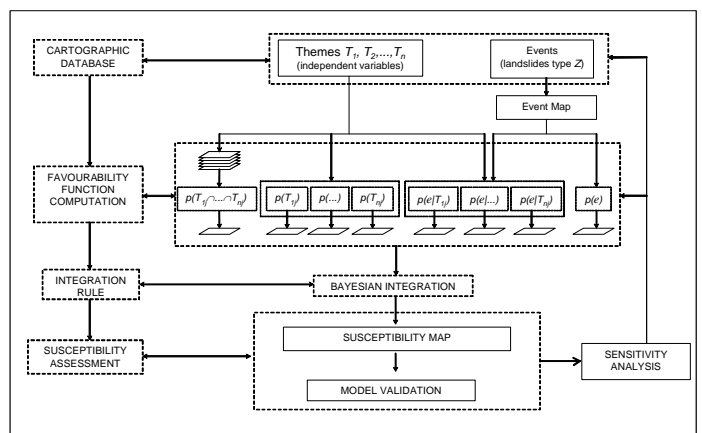


Figure 2. Methodology for landslide susceptibility assessment and validation.

The calculation of *a priori* and conditional probabilities was the first step in the cartographic data integration. These probabilities were estimated by

overlapping the landslide map to those maps representing each landslide predisposing factor, and using the following equations:

(i) *a priori* probability of landslide occurrence

$$Pp_s = \left(\frac{S_{area}}{T_{area}} \right) \quad (1)$$

where S_{area} = landslide area within the test site; T_{area} = total area of test site.

(ii) *a priori* probability of occurrence of a class x belonging to the predisposing factor T

$$Pp_c = \left(\frac{X_{area}}{T_{area}} \right) \quad (2)$$

where X_{area} = area of class x from the predisposing factor T .

(iii) conditional probability of landslide occurrence in the class x from the predisposing factor T

$$Cp = 1 - \left(1 - \frac{1}{X_{area}} \right)^{S_x} \quad (3)$$

where S_x = landslide area within class x from the predisposing factor T .

Scores obtained by the application of equation (3) for each class of each considered landslide predisposing factor are interpreted as favorability values, or landslide susceptibility indicators.

The probability of landslide occurrence given n landslide predisposing maps is obtained using the conditional probability integration rule through the next expression (Chung & Fabbri 1999, Zêzere et al. 2004):

$$P = \frac{(P_{pT1} * P_{pT2} * ... * P_{pTn})(C_{pT1} * C_{pT2} * ... * C_{pTn})}{Pp_{slide}^{Tn-1} * (T1 * T2 * ... * Tn)} \quad (4)$$

where $T1, T2, ..., Tn$ = set of landslide predisposing factors; P_p = *a priori* probability of occurrence of a class x from the predisposing factor T ; Pp_{slide} = *a priori* probability of landslide occurrence; C_p = conditional probability of occurrence of a landslide in the class x from the predisposing factor T .

The equation (4) was applied on a 5 m grid cell structure that is reasonably conform to the detail and resolution of the cartographic database (Zêzere et al. 2007). The obtained results (a *score* for each pixel of the study area) range between 0 and 1 and measure the susceptibility (or spatial probability) of occurrence of future shallow translational slides.

The prediction model performance was assessed through the computation of success rate curves (Fabbri et al. 2002). These curves were constructed by crossing the distribution of the total set of landslides used to generate the susceptibility model with the prediction results, after sorting in descend-

ing order the susceptibility values corresponding to each pixel. Additionally, we compute the 'Area Under the Curve' (AUC) for each success rate curve, in order to quantify the model performance, and to allow the objective comparison among different success rate curves. The AUC values range from 0 to 1, and the quality of the prediction model increases with the increase of the AUC value.

The landslide susceptibility model was applied, in a first step, to each landslide predisposing factor considered individually, and seven prediction rate curves were constructed. The AUC values corresponding to these curves were used to rank variables. In the next step, the landslide susceptibility model was performed using groups of 2, 3, 4, 5, 6 and 7 predisposing factors that were selected according to the above mentioned AUC-based ranking.

Lastly, the models corresponding to the best variable combinations are tested by the computation of prediction rate curves (Chung & Fabbri 2005), based on the temporal partition of the original landslide data base in two parts: prediction set (used to develop the landslide prediction model) and validation set (used for the independent validation of the predicted results).

4 RESULTS AND DISCUSSION

As previously referred, one of the main goals of the sensitivity analysis was to assess the weight of different landslide predisposing factors within a statistically-based landslide susceptibility model. Therefore, Figure 3 and Table 2 illustrate, respectively, the success rate curves and the corresponding AUC obtained by applying the predictive model using separately each one of the slope instability predisposing factors. The obtained results demonstrate that the considered independent variables do not correlate in the same way with landslide distribution (AUC ranging from 0.67 to 0.80). Moreover, according to AUC records, 'slope angle' and 'geomorphological units' are the variables more able to predict the future occurrence of shallow translational slides.

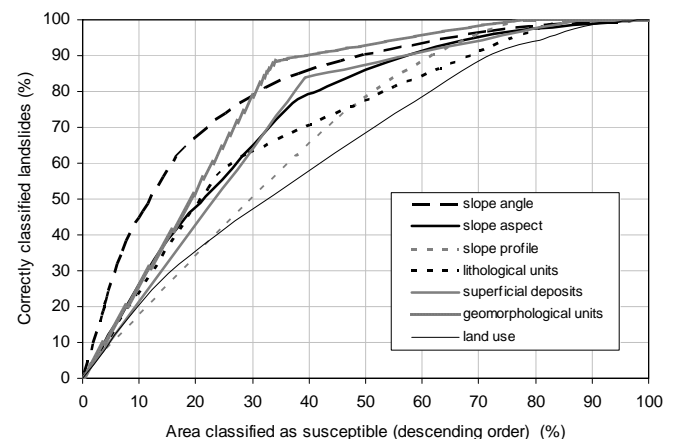


Figure 3. Success rate curves corresponding to individual landslide predisposing factors.

Table 2. Hierarchy of predisposing factors for shallow translational slides occurrence, according to success rate curves and AUC (Area Under the Curve).

Rank	Variable ID	Variable name	AUC
1	A	Slope angle	0.802
2	F	Geomorphological units	0.786
3	B	Slope aspect	0.738
4	E	Superficial deposits	0.732
5	D	Lithological units	0.703
6	C	Transversal slope profile	0.671
7	G	Land use	0.633

The variable ranking summarized in Table 2 was used to define the conjugation of landslide predisposing factors that support the next landslide susceptibility models (i.e. models running with 2 variables, 3 variables, 4 variables, 5 variables, 6 variables and 7 variables). Figure 4 illustrates the success rate curves of these landslide prediction models, and Table 3 summarizes the corresponding ‘Area Under the Curve’. Additionally, Figure 5 illustrates the variation on models prediction capability according to the number of variables within the model, for some standard areas of maximum landslide susceptibility (corresponding to 5%, 10%, 20%, 30% and 40% of the total study area).

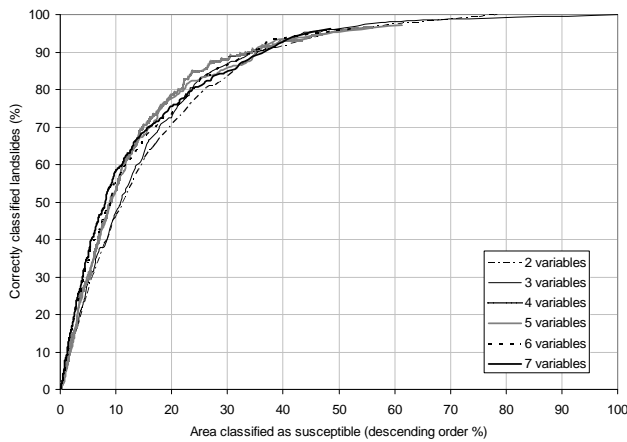


Figure 4. Success rate curves corresponding to landslide susceptibility models obtained using from 2 to 7 predisposing factors (2 variables = variable Id: A+F; 3 variables = variable Id: A+F+B; 4 variables = variable Id: A+F+B+E; 5 variables = variable Id: A+F+B+E+D; 6 variables = variable Id: A+F+B+E+D+C; 7 variables = total variable set).

Table 3. Area under the curve (AUC) of success rate curves corresponding to landslide susceptibility models obtained using from 2 to 7 predisposing factors.

Variables in the model	AUC
2 variables (variable Id: A+F)	0.839
3 variables (variable Id: A+F+B)	0.847
4 variables (variable Id: A+F+B+E)	0.857
5 variables (variable Id: A+F+B+E+D)	0.852
6 variables (variable Id: A+F+B+E+D+C)	0.854
7 variables (variable Id: A+F+B+E+D+C+G)	0.857

The analysis of Figures 4 and 5 and Table 3 allow concluding that:

- the quality of landslide prediction models demonstrates a slight tendency to improve with the increment on the number of variables within the model, as it is shown by the AUC values (Table 3). This is particularly true when we consider the top 5% and 10% of the total area classified as more susceptible to slope instability (Fig. 5);
- if we consider the 30% and 40% of the total area classified as more susceptible, the predicted results tend to stabilize (Fig. 5), with maximum variations of 4% on success rate curves. These features demonstrate the low sensitivity of landslide prediction models to the increasing number of landslide predisposing factors;
- the introduction of more variables in the landslide prediction model, does not generate necessarily better results in success rates. For instance, the model produced using 4 variables generate better prediction results, when compared with those obtained using 5 and 6 variables, as it is confirmed by the corresponding AUC values (Table 3).
- it is possible to predict the future spatial occurrence of shallow translational slides in the study area with very satisfactory results, based on a restricted number of landslide predisposing factors. For instance, the susceptibility model produced with 4 variables (slope angle, geomorphological units, slope aspect and superficial deposits) shows results very similar to those obtained using the total set of variables, as it is confirmed by the shape of success rate curves (Fig. 4) and the corresponding AUC values (Table 3). Moreover, results obtained with the above mentioned 4 landslide predisposing factors are even better than those obtained with the complete set of variables when we isolate for analysis the 20% and 30% of area defined as more susceptible (Fig. 5).

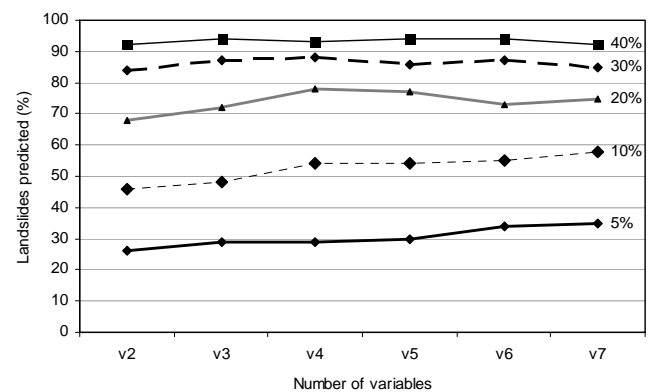


Figure 5. Variation on the predictive power of landslide susceptibility models according to the number of variables in the model (higher susceptibility scores corresponding to 5%, 10%, 20%, 30% and 40% of the total area were selected for comparison).

Figure 6 and Figure 7 show the susceptibility maps to shallow translational slides occurrence in the study area, based on 4 predisposing factors

(slope angle, geomorphological units, slope aspect and superficial deposits) and on the total set of predisposing factors, respectively. In order to allow map comparison, we define 6 landslide susceptibility classes, which were generated in the same way for both maps as % of the total area, after sorting in descending order the susceptibility values corresponding to each pixel. Table 4 summarizes the spatial probabilities computed for each 1% of the total study area, for landslide susceptibility classes represented in Figures 6 and 7. The obtained results are very similar, despite the higher probability corresponding to the first susceptibility class in the map based on the complete set of predisposing factors (Table 4).

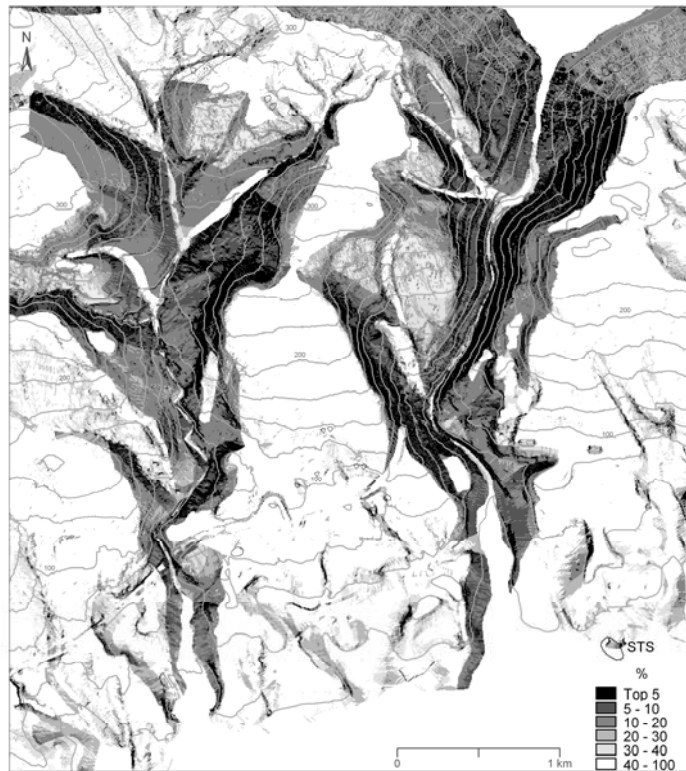


Figure 6. Shallow translational slides susceptibility map based on 4 predisposing factors (slope angle, geomorphological units, slope aspect and superficial deposits). Susceptibility classes were defined as % of the total area, after sorting in descending order the susceptibility values corresponding to each pixel. STS = shallow translational slides.

The comparison between Figure 6 and Figure 7 allows concluding that the obtained spatial patterns of landslide susceptibility are very similar. This fact is more evident for the high susceptibility classes, which is in accordance with the success rate data. The main difference found between landslide susceptibility maps is the homogeneity level of representation of susceptibility classes, which is higher in the prediction model reported in Figure 6. This difference is explained by the different number of ‘unique conditions’ (particular terrain combinations) presented within landslide prediction models. These ‘unique conditions’ result from the crossing among the set of predisposing maps considered in each

model. Therefore, the prediction model based on 7 variables generates 15,636 unique conditions that contribute to a less regular pattern of susceptibility distribution (Fig. 7). On the other hand, the prediction model supported by 4 variables generates only 234 unique conditions, which justify the higher spatial homogeneity of susceptibility classes (Fig. 6).

Finally, the prediction capability of landslide susceptibility models based on 4 and 7 landslide predisposing factors was assessed through the construction of prediction rate curves. These curves are shown in Figure 8 and were constructed by dividing the original landslide data base in two groups using a temporal criterion: the landslide prediction group (46 shallow translational slides occurred prior to 1980 that were used to develop a new landslide prediction model); and the landslide validation group (54 shallow translational slides occurred after 1980 that were used for the independent validation of the predicted results).

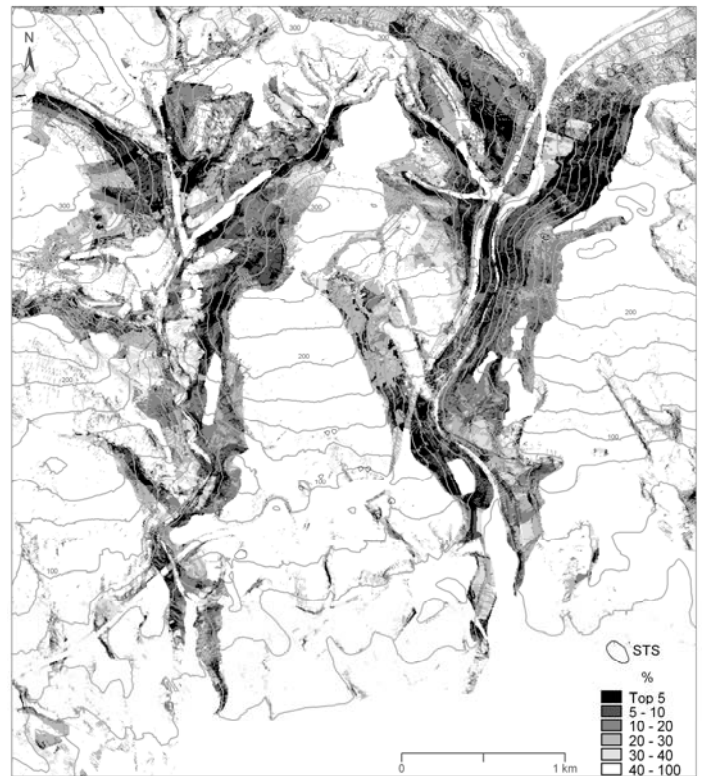


Figure 7. Shallow translational slides susceptibility map based on the complete (7) set of predisposing factors. Susceptibility classes were defined as % of the total area, after sorting in descending order the susceptibility values corresponding to each pixel. STS = shallow translational slides.

Table 4. Estimated probability (%) for landslide susceptibility classes represented in Figures 6 and 7, per 1% of total area.

	Top 5	5-10	10-20	20-30	30-40	40-100
Figure 6	5.9	4.8	2.5	1.0	0.5	0.1
Figure 7	7.0	4.6	1.7	0.9	0.9	0.1

As it was expected, both prediction rate curves are below the corresponding success rate, which were produced using the same landslide data set for

modeling and validation. Anyway, the prediction rate curves show fairly acceptable results for both models performed with 4 and 7 variables. For instance, 47% to 50% of landslides occurred after 1980 are within the 10% of the total area classified as more susceptible (Fig. 8). These features grow up to 72% if we consider the 20% of the total area classified as more susceptible by both prediction models (Fig. 8).

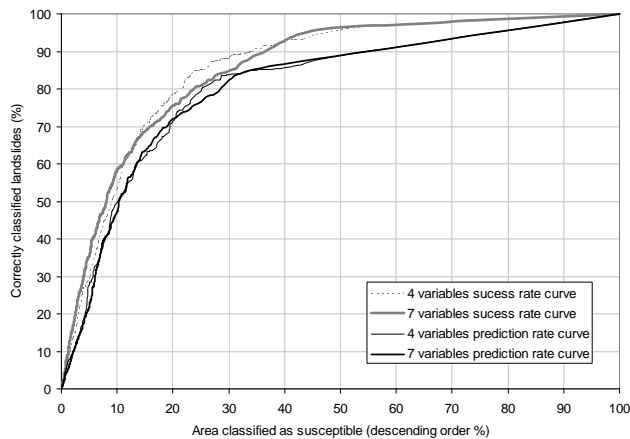


Figure 8. Prediction rate curves of landslide susceptibility models obtained using 4 predisposing factors (slope angle, geomorphological units, slope aspect and superficial deposits) and 7 predisposing factors (corresponding success rate curves are also plotted for comparison).

The prediction rate curves corresponding to landslide susceptibility models produced using 4 and 7 landslide predisposing factors are very similar (Fig. 8) allowing to conclude that prediction performance is equivalent for both models. This fact is also confirmed by the corresponding AUC: 0.809 for the model based on 4 variables; 0.808 for the model based on 7 variables.

5 CONCLUSION

The relationship between the number of predisposing factors within a statistically-based landslide prediction model and the quality of predicted results is not linear. The results obtained here prove that the introduction of additional variables into a landslide prediction model does not generate necessarily better success rates. Moreover, as it was shown in the present analysis, it is possible to obtain an accurate landslide susceptibility map using a limited number of instability predisposing factors in the prediction model (e.g. slope angle, geomorphological units, slope aspect and superficial deposits). However, these “key variables” cannot be extrapolated for other types of landslides within the test site or for other study areas. Therefore, a prudent approach to landslide susceptibility assessment implies, on a first step, the use of a set of coherent and logical landslide predisposing factors as large as possible.

On a second step, the landslide predictive models can be simplified, with minor losses, by removing those variables that prove to be irrelevant to the prediction performance of the susceptibility model.

ACKNOWLEDGEMENTS

The research of R.A.C. Garcia and S.C. Oliveira was supported by the Portuguese Foundation for Science and Technology of the Portuguese Ministry of Science, Technology and Higher Education.

REFERENCES

- Carrara, A.; Guzzetti, F.; Cardinali, M. & Reichenbach, P. 1999. Use of GIS technology in the prediction and monitoring of landslide hazard. *Natural Hazards* 20: 117-135.
- Chung, C.F. & Fabbri, A. 1993. The representation of geo-science information for data integration. *Nonrenewable Resources* 2 (2): 122-138.
- Chung, C.F. & Fabbri, A. 1999. Probabilistic prediction models for landslide hazard mapping. *Photogrammetric Engineering and Remote Sensing* 65 (12): 1389-1399.
- Chung, C.F. & Fabbri, A. 2005. Systematic procedures of landslide-hazard mapping for risk assessment using spatial prediction models. In Glade, T., Anderson, M.G., Crozier, M.J. (eds.), *Landslide Hazard and Risk*: 139-174. Chichester: Wiley.
- Fabbri, A.; Chung, C.F.; Napolitano, P.; Remondo, J. & Zêzere, J.L. 2002. Prediction rate functions of landslide susceptibility applied in the Iberian Peninsula. In Brebbia C.A. (ed.), *Risk Analysis III, Series. Management Information Systems* vol. 5: 703-718. Southampton, Boston: WIT Press.
- Guzzetti, F.; Carrara, A.; Cardinali, M. & Reichenbach, P. 1999. Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. *Geomorphology* 31: 181-216.
- Guzzetti, F.; Reichenbach, P.; Cardinali, M.; Galli, M. & Ardizzone, F. 2005. Probabilistic landslide hazard assessment at the basin scale. *Geomorphology* 72: 272-299.
- Reis, E.; Zêzere, J.L.; Vieira, G.T. & Rodrigues, M.L. 2003. Integração de dados espaciais em SIG para avaliação da susceptibilidade à ocorrência de deslizamentos. *Finisterra XXXVIII* (76): 3-34.
- Zêzere, J.L.; Garcia, R.A.C.; Oliveira, S.C. & Reis, E. 2007. Probabilistic landslide risk analysis considering direct costs in the area north of Lisbon (Portugal). *Geomorphology*. DOI: 10.1016/j.geomorph.2006.10.040 (in press)
- Zêzere, J.L.; Reis, E.; Garcia, R.; Oliveira, S.; Rodrigues, M.L.; Vieira, G. & Ferreira, A.B. 2004. Integration of spatial and temporal data for the definition of different landslide hazard scenarios in the area north of Lisbon (Portugal). *Natural Hazards and Earth System Sciences* 4: 133-146.
- Zêzere, J.L. & Rodrigues, M.L. 2002. Rainfall Thresholds for Landsliding in Lisbon Area (Portugal). In: Rybar, Stemberk, Wagner (eds.), *Landslides*: 333-338. Lisse: Swets & Zeitlinger.
- Zêzere, J.L.; Trigo, R. & Trigo, I. 2005. Shallow and deep landslides induced by rainfall in the Lisbon region (Portugal): assessment of relationships with the North Atlantic Oscillation. *Natural Hazards and Earth System Sciences* 5: 331-344.