Rend. Online Soc. Geol. It., Vol. 21 (2012), pp. 401-403, 4figs, 1 tab.

Shallow-landslide susceptibility evaluation by means of Logistic Regression. The Costa Viola case study (Southern Calabria, Italia)

R. GRECO (1), O.G. TERRANOVA (1), A.D. PELLEGRINO (2), S.L. GARIANO (1), P. IAQUINTA (1), G. MANDAGLIO (3), A. LA TORRE (4), G. IOVINE (1)

Key words: *Shallow-landslide, Susceptibility, Logistic Regression.*

INTRODUCTION

Shallow landslides are among the most destructive and dangerous types of slope movements for people and infrastructures. The “Costa Viola” study area (Southern Calabria), especially in the sector between Bagnara Calabra and Scilla, has repeatedly been affected by slope instability events in the past, mainly related to shallow landslides (debris slides and debris flows) (PELLEGRINO and BORRELLI, 2007).

Evaluation of landslide susceptibility, the first step in regional hazard management, can be achieved by means of multivariate statistical analyses (GUZZETTI *et alii*, 1999). Logistic Regression (LR) is a type of multivariate analysis widely utilized (cf. among the most recent studies: GRECO *et alii*, 2007; SORRISO-VALVO *et alii*, 2009; ROSSI *et alii*, 2010); it allows estimating the presence/absence of a phenomenon, represented by the dependent variable *y*, in terms of probability P(y), on the basis of linear statistical relationships with a set of independent variables. The adopted LR procedure consists of four steps: i) variable parameterization, ii) sampling, iii) fitting, and iv) application.

An attempt of improving preliminary results, recently presented at the Second World Landslide Forum (IOVINE *et alii*, 2011), has been performed by considering other territorial variables, the use of different variable parameterization, and by processing data through a specific statistical package.

STUDY AREA

The study area (82.2 km^2) is located in the “Costa Viola” mountain ridge, between Bagnara Calabra and Scilla (Fig. 1). It is characterized by Palaeozoic metamorphic and crystalline bedrocks, covered by Upper Miocene to Holocene sedimentary deposits. At the base of the mountain ridge, a NE-SW trending fault marks the transition between the basement and the sedimentary terrains of the coastal plain.

Morphologically, the area is characterized by steep slopes, cut by deep canyons (the longest thereof are T. Favazzina and T. Sfalassà); a set of marine terraces can be recognized between 100 and 600 m a.s.l. The study area is crossed by relevant infrastructures (railway, highway A3, State road SS.18) (cf. Fig. 1); the narrow coastal plain is densely urbanized.

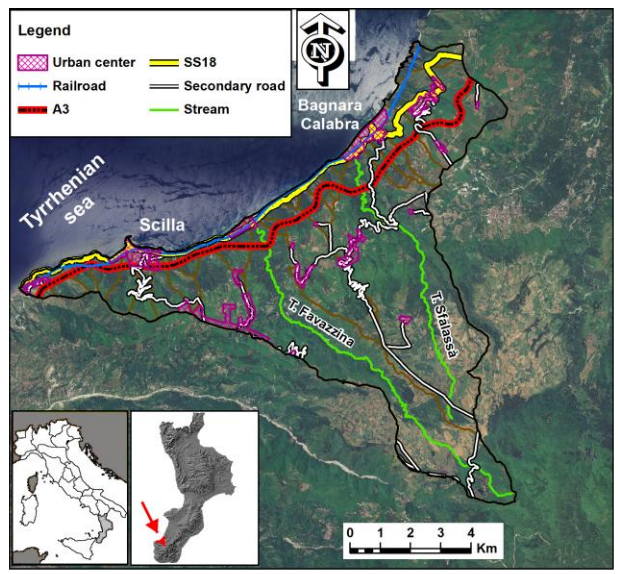

Fig. 1 – Location and main infrastructures of the study area

METHOD

In LR, the presence/absence of a landslide can be expressed in terms of a dichotomic dependent variable, whose probability of being true is determined on the basis of statistical relationships with a set of independent territorial variables. The probability of occurrence of the dependent variable, P_y, can be calculated as follows:

$$P_y = 1 - \left[1 + exp\left(-\sum a_i X_i\right)\right] \qquad [1]$$

(1) CNR-IRPI - U.O.S. di Cosenza, via Cavour, 6 – 87036 Rende (CS) Italia – e-mail: g.iovine@irpi.cnr.it

(2) ABR Calabria, Catanzaro (CZ), Italia.

(3) Università degli Studi Mediterranea di Reggio Calabria, Italia.

(4) Provincia di Reggio Calabria, Italia.

in which a_i and X_i are the regression coefficients (i.e., the variables weights), and the independent variables, respectively.

The adopted LR procedure, first proposed by GRECO *et alii* (2007), consists of four steps: i) variable parameterisation, performed by considering landslide incidence in given portions of the study area, named the "sampling zones"; ii) sampling, in which a fitting data set is extracted from buffer sampling zones, obtained by generating a buffer around each landslide source; iii) model fitting, iteratively performed by testing all the possible values for the weights of the independent variables, until the quadratic mean residuals are minimized; iv) model application, in which the logistic function is applied to the whole study area. A grid of values, ranging between 0 and 1, is obtained as output. The forecast capability can be evaluated by means of the ROC analysis (Receiver Operating Characteristics Analysis, Hosmer and Lemeshow, 1989).

The territorial distribution of shallow landslides (i.e. the dependent variable) has been obtained by interpreting a set of aerial photographs (1954-1955). In Fig. 2, the 181 landslides employed as training set during the phase of calibration are shown.

Independent variable maps have been obtained through field survey, from available published maps, and by processing other variables (derived from the DEM).

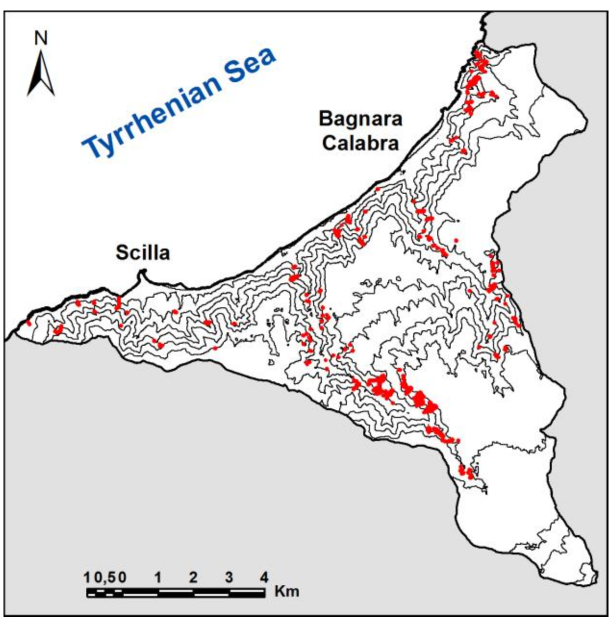

Fig. 2 – Landslide inventory maps of the study area.

The data set of independent variables includes the nine causal factors (Fig. 3a-h): lithology (LU), land use (LUS), soil type (ST), elevation (ELEV), slope angle (SLO), aspect (ASP), curvatures set along the direction of maximum slope (ACUR), curvatures set orthogonally to the direction of maximum slope (DCUR), topographic wetness index (TWI).

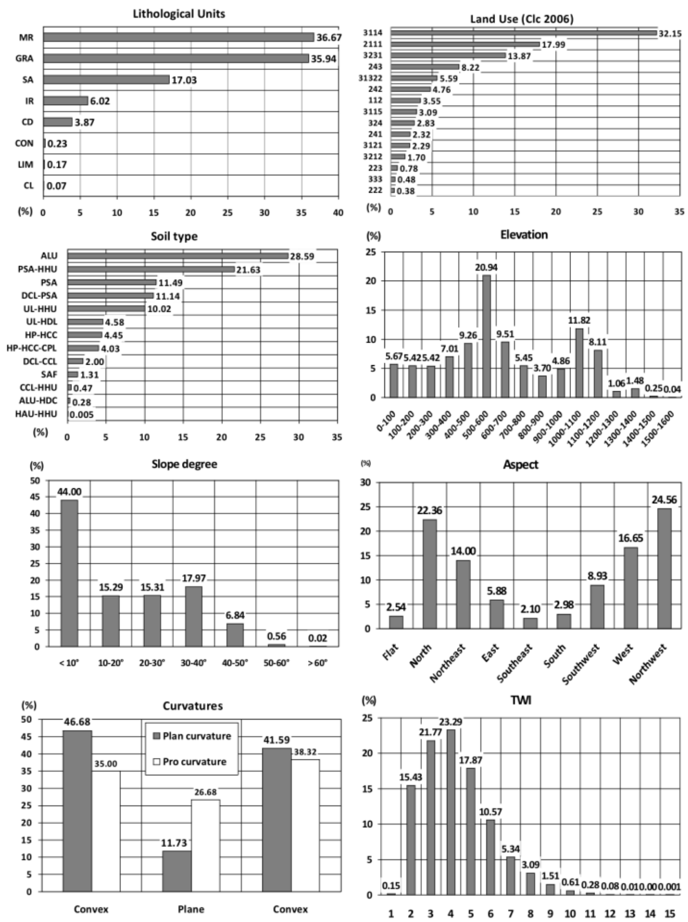

Fig. 3 – Distribution of independent territorial variables for the study area: Keys: a) Lithological Units: CL = Clay; LIM = Limestone; CON = Conglomerate; CD = Colluviums and debris; IR = Igneous Rock; SA = Sand; GRA = Gravel; MR = Metamorphic Rock; b) Land Use Categories: 112 = Discontinuous urban fabric; 2111 = Intensive cultivations; 222 = Fruit trees and berries plantations; 223 = Olive groves; 241 = Annual crops associated with permanent crops; 242 = Complex cultivation patterns; 243 = Land principally occupied by agriculture, with significant areas of natural vegetation; 3114 = Chestnut forest; 3115 = Beech forest; 3121 = Coniferous forest; 31322 = Mixed forest; 3212 = Discontinuous grassland; 3231 = High "Macchia"; 324 = Transitional woodland-shub; 333 = Sparsely vegetated areas; c) Soil type: ALU = Artenic-Leptic Ubrisol; CCL = Cutanic Chromic Luvisols; CPL = Chromi Profondic Luvisols; DCL = Dystri-Cutanic Luvisol; HAU = Humi-Arenic Umbrisols; HCC = Haspli-Calcaric Cambisols; HDC = Hapli-Dystic Cambisols; HDL = Hapli-Dystric Leptosols; HHU = Hapli-Humic Umbrisols; HP = Haplic Phaeozems; PSA = Pachi-Silic Andosols; SAF = Skeletic-Arenic Fluvisols; UL = Umbrihumic Leptosols.

PRELIMINARY RESULTS AND IMPROVEMENTS

In Table 1, regression weights of each independent variable, RMSE and Chi-square values obtained in the model fitting are shown. Through the logistic function [1], the probability of existence has been evaluated for the whole study area. Finally, the probability has been re-classified into 5 classes: Null ($P(y) \leq 5\%$); Low ($5\% < P(y) \leq 25\%$); Medium ($25\% < P(y) \leq 55\%$); High ($50\% < P(y) \leq 75\%$); Very high ($P(y) > 75\%$) (Fig. 4).

coefficient	variable	value
a_0	intercept	-11.552
a_1	LU	0.028
a_2	LUS	0.013
a_3	ST	0.000
a_4	ELEV	0.018
a_5	SLO	0.025
a_6	ASP	0.032
a_7	ACUR	0.027
a_8	DCUR	0.043
a_9	TWI	0.037
RMSE		0.485
Chi-square		475.990

Tab. 1 – Regression weights, RMSE and Chi-square values.

As shown in Figure 4, the sectors most threatened by shallow-landslides are located by the Tyrrhenian coast, and along the steep flanks of the Sfalassà and Favazzina torrents. Obtained weights (Tab. 1) suggest that down-slope curvature, topographic wetness index, aspect and lithology play a primary role in favouring slope instability.

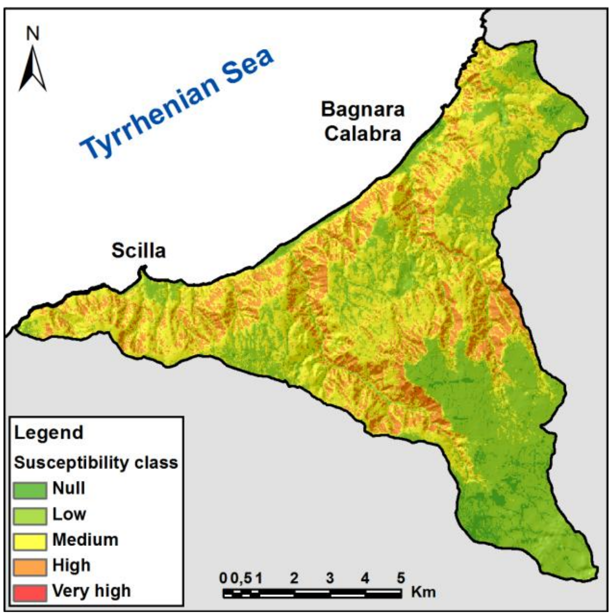

Fig. 4 – Susceptibility map of Costa Viola study area.

In order to check the performance of the LR models, a ROC analysis has been performed. The obtained ROC-curve (Fig. 5) indicates a percentage of 85.6% of correctly classified cells, confirming the goodness of classification for the “Costa Viola” study area, where the LR sampling procedure satisfactorily described the probability of presence of mass movements.

The next step of the study, still in progress, is an attempt of improving the cited preliminary results, through: i) introducing other territorial variables (like the distance to fault, the soil thickness, the weathering grade); ii) using different variable parameterization (dummy approach); iii) processing data through a specific statistical software (like IBM SPSS Statistics 19).

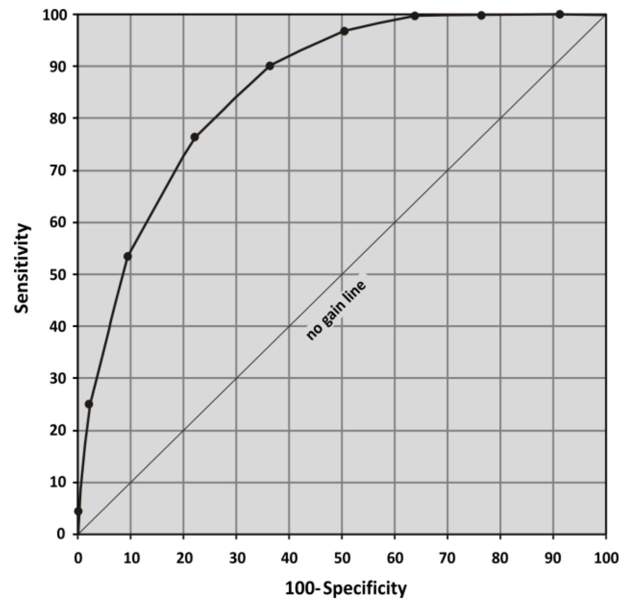

Fig. 5 – Obtained ROC curve.

REFERENCES

GRECO R., SORRISO-VALVO G.M. & CATALANO E. (2007) - *Logistic Regression analysis in the evaluation of mass-movement susceptibility: the Aspromonte case study, Calabria, Italy*. Engineering Geology, **89**, 47-66.

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.

IOVINE G.G.R., GRECO R., GARIANO S.L., IAQUINTA P., PELLEGRINO A. & TERRANOVA O.G. (2011) - *Shallow-landslide susceptibility in the Costa Viola mountain ridge (Italia).* Proceedings of the Second World Landslide Forum – 3-7 October 2011, Rome.

PELLEGRINO A. & BORRELLI S., (2007) - *Analisi del dissesto da frana in Calabria. Rapporto sulle frane in Italia. Il progetto IFFI – Metodologia, risultati e rapporti regionali.* Rapporti 78/2007. Trigila A (ed). APAT, Roma, 681pp.

ROSSI M., GUZZETTI F., REICHENBACH P., MONDINI A. C. & PERUCCACCI S. (2010) - *Optimal landslide susceptibility zonation based on multiple forecast*. Geomorphology, **114**, 129-142.

SORRISO-VALVO G.M., GRECO R. & CATALANO E. (2009) - *Spatial prediction of regional scale mass movement using the Logistic Regression analysis and GIS – Calabria, Italy*. Israel Journal of Earth Sciences, **57**, 263-268.