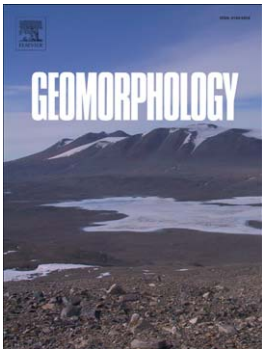


Exploring the effect of absence selection on landslide susceptibility models: A case study in Sicily, Italy

Christian Conoscenti, Edoardo Rotigliano, Mariaelena Cama, Nathalie Almaru Caraballo-Arias, Luigi Lombardo, Valerio Agnesi

PII: S0169-555X(16)30096-4
DOI: doi: 10.1016/j.geomorph.2016.03.006
Reference: GEOMOR 5538

To appear in: *Geomorphology*

Received date: 2 August 2015
Revised date: 24 November 2015
Accepted date: 10 March 2016

Please cite this article as: Conoscenti, Christian, Rotigliano, Edoardo, Cama, Mariaelena, Caraballo-Arias, Nathalie Almaru, Lombardo, Luigi, Agnesi, Valerio, Exploring the effect of absence selection on landslide susceptibility models: A case study in Sicily, Italy, *Geomorphology* (2016), doi: 10.1016/j.geomorph.2016.03.006

Exploring the effect of absence selection on landslide susceptibility models: A case study in Sicily, Italy

Christian Conoscenti[a,*], Edoardo Rotigliano[a], Mariaelena Cama[a], Nathalie Almaru Caraballo-Arias[b], Luigi Lombardo[a], Valerio Agnesi[a]

[a] Department of Earth and Marine Sciences (DISTEM), University of Palermo, Via Archirafi 22, 90123 Palermo, Italy

[b] Department of Agricultural and Forestry Sciences, University of Palermo, Viale delle Scienze, 90128, Palermo, Italy

* Corresponding author (C. Conoscenti): Tel: +3909123864670; Fax: +39 0916169908; E-mail address: christian.conoscenti@unipa.it

Abstract

A statistical approach was employed to model the spatial distribution of rainfall-triggered landslides in two areas in Sicily (Italy) that occurred during the winter of 2004–2005. The investigated areas are located within the Belice River basin and extend for 38.5 and 10.3 km^2, respectively. A landslide inventory was established for both areas using two Google Earth images taken on October 25th 2004 and on March 18th 2005, to map slope failures activated or reactivated during this interval. Geographic Information Systems (GIS) were used to prepare 5 m grids of the dependent variables (absence/presence of landslide) and independent variables (lithology and 13 DEM-derivatives). Multivariate Adaptive Regression Splines (MARS) were applied to model landslide susceptibility whereas receiver operating characteristic (ROC) curves and the area under the ROC curve (*AUC*) were used to evaluate model performance. To evaluate the robustness of the whole procedure, we prepared 10 different samples of positive (landslide presence) and negative (landslide absence) cases for each area. Absences were selected through two different methods: (i) extraction from randomly distributed circles with a diameter corresponding to the mean width of the landslide source areas; and (ii) selection as randomly distributed individual grid cells. A comparison was also made between the predictive performances of models including and not including the lithology parameter.

The models trained and tested on the same area demonstrated excellent to outstanding fit (*AUC* > 0.8). On the other hand, predictive skill decreases when measured outside the calibration area, although most of the landslides occur where susceptibility is high and the overall model performance is acceptable (*AUC* > 0.7). The results also showed that the accuracy of the landslide susceptibility models is higher when lithology is included in the statistical analysis. Models whose absences were selected using random circles showed a significantly better performance when learning and validation samples were extracted from the same area; whereas, conversely, no significant difference was observed when testing the models outside the training area.

Keywords: Landslide susceptibility; Google Earth; Geographic Information Systems (GIS); Multivariate Adaptive Regression Splines (MARS).

1. Introduction

Landslide susceptibility is defined as the proneness of a terrain unit to generate landslides (Brabb, 1984; Carrara et al., 1995; Guzzetti et al., 1999). A map of landslide susceptibility expresses, typically in relative terms, the spatial likelihood of landslide occurrence within a given territory. As the occurrence of slope failures may have severe economic and social consequences, landslide susceptibility maps can assist land managers and policy makers in implementing land-use strategies to reduce landslide hazard.

Landslide susceptibility may be assessed using both direct methods based on expert geomorphological analysis and indirect methods relying on deterministic or stochastic approaches. Over the last decades, the statistical approach to landslide susceptibility modeling has become very popular due to the increasing availability of low cost high-resolution data, and the development of open-source statistical software and Geographical Information Systems (GIS). This approach is based on the assumption that new landslides are more likely to occur under environmental conditions similar to those that led to past slope failures (Carrara et al., 1995; Guzzetti et al., 1999; van Westen et al., 2005, 2008). The approach requires a landslide inventory and a set of environmental attributes related to the occurrence of slope failures. Landslide inventories are usually made by integrating field surveys with analyses of high quality aerial/satellite images. Presence or absence of landslides within a mapping unit (e.g., grid cell, slope unit, and terrain unit) represents the dependent variable, which is predicted by an ensemble of independent environmental variables. The variables are proxies of the main landslide triggering factors and are selected according to their relevance to slope stability and the quality and resolution of available data. Statistical analysis of landslide susceptibility exploits either bivariate modeling techniques (e.g., Agnesi et al., 1982; Carrara et al., 1995; Clerici et al., 2002; Vergari et al., 2011; Rotigliano et al., 2012) or multivariate ones (e.g., Van Den Eeckhaut et al., 2006; Atkinson and Massari, 2011; Conforti et al., 2014; Cama et al., 2015; Goetz et al., 2015). Comprehensive reviews of statistical

models employed in the field of landslide susceptibility modeling can be found in Aleotti and Chowdhury (1999), Guzzetti et al. (1999) and Brenning (2005).

Most of the statistical models employed to predict landslide spatial distribution are fitted to data sets with both positive (landslide presence) and negative (landslide absence) cases. Positives are often sampled from subsets of grid cells containing one to all cells within each landslide, whereas negatives are typically randomly selected as individual pixels outside the landslide areas. Then, landslide susceptibility models are calibrated and validated exploiting different samples of data, but typically extracted from the same study area, performing a random partition of positives and negatives (Chung and Fabbri, 2003). Relatively few landslide susceptibility studies have attempted a validation with independent data from areas outside those used to calibrate the models (e.g., Von Ruette et al., 2011; Pradhan et al., 2010; Costanzo et al., 2012a; Lombardo et al., 2014).

In this experiment we employed Multivariate Adaptive Regression Splines (MARS; Friedman, 1991) to model the spatial distribution of landslides that were triggered in two study areas of Sicily (Italy) by rainfall during the 2004–2005 winter season. A landslide inventory was established for both areas through the analysis of two Google Earth images, dated October 25th, 2004 and March 18th, 2005, by mapping slope failures triggered or reactivated during this time period. A limited number of studies have exploited Google Earth images to prepare landslide inventories (e.g., Costanzo et al., 2012a,b; Schicker and Moon, 2012; Van Den Eeckhaut et al., 2012; Borrelli et al., 2014; Zhang et al., 2015) and, as far as we know, none produced their susceptibility models without incorporating information from field surveys and/or other sources of data. Indeed, conventional methods to prepare landslide inventory rely mainly on geomorphological field mapping and on the interpretation of stereoscopic aerial photographs (Guzzetti et al., 2012). Conversely, in this study, the 3D view provided by the Google Earth software was the only tool used for landslide detection and mapping. This allowed us to test whether effective landslide susceptibility models may be prepared without field mapping. In this experiment, we used one area to both calibrate and validate landslide susceptibility models whereas the other area was only used to assess the predictive skill of

the models trained for the first area. To test the robustness of the procedure, 10 training and 10 test samples were extracted from the first area, and 10 validation subsets were identified in the second area. They were prepared by adopting different strategies to select landslide absences: (i) extraction from randomly distributed circles having a diameter corresponding to the mean width of the identified landslide source areas; and (ii) selection as randomly distributed individual grid cells. Moreover, we prepared models that both included and did not include lithology as a predictor variable. The main objectives of this experiment were to: (i) evaluate whether landslide inventories based on Google Earth images as their only data source can be used to prepare reliable landslide susceptibility models; (ii) explore how the performance of landslide susceptibility models is affected by changing the method to sample landslide absences; (iii) assess the accuracy of landslide predictions outside the area where the models were calibrated; and (iv) evaluate the importance of lithology as a predictor of landslide distribution.

2. Materials and methods

2.1. Study areas

Two study areas were selected for this experiment. Both areas are located within the catchment of the Belice River (Fig. 1), one of Sicily's main river basins. The two areas, hereafter referred to as AREA1 and AREA2 (Fig. 2), extend for 38.5 and 10.3 km^2, respectively. Their altitudes are 218–519 m a.s.l. (mean = 371.6 m and std. dev. = 68.9 m) and 317–714 m a.s.l. (481.2 and 78.8 m), respectively. The slope gradient of AREA1 (mean = 9.5° and std. dev. = 6.0°) tends to be slightly lower than that of AREA2 (10.8° and 5.6°).

The study area's climate is Mediterranean, with hot and dry summers and mild and wet winters. According to the rainfall data from the meteorological station in Corleone (588 m a.s.l.), the average annual rainfall is 643.3 mm. Precipitation occurs mainly during the autumn–winter semester, with peaks in December (91.3 mm) and January (82.5 mm) (Fig. 3).

The two study areas are mainly characterized by hilly landscapes, where slope and channel processes prevail. However, despite the very small distance separating them (around 3 km), their geological and geomorphological settings are different. AREA1 is mainly underlain by clays of the Late Miocene Terravecchia Formation (31% of the total extent) and by marls and sandstones of the Late Pliocene–Early Pleistocene Marnoso-Arenacea Formation (26%) (Table 1 and Fig. 4). The former lithology is dominant in the northern and western parts, which are characterized by gentle slopes, drained by a relatively wide and shallow valley. The latter prevails in the SW sector, where, due to the outcropping of harder rocks, the land surface is more rugged, with deeper valleys and steeper slopes. AREA2 corresponds to the lowest part of a broad valley, running approximately E–W. This area is mainly underlain by marls of the Middle–Late Miocene Marne di San Cipirello Formation and the Late Oligocene–Early Miocene Marne di Cardellia Formation (43% of the total extent). The flanks of the valley are gentle and partially covered by ancient landslides, which extend for 28% of AREA2.

Intense water erosion and gravitational processes affect both study areas. Landslides generally consist of earth-flows triggered by rainfall during autumn and winter. These phenomena cause damage to infrastructure, such as roads and walls, and affect agricultural lands, which are the dominant land cover (Corine Land Cover 2006 by the European Environment Agency, 2010) and the main economic activity of the study areas.

2.2. Landslide inventories

Landslide inventories of both study areas were prepared by visual interpretation of two satellite images available from Google Earth, dated October 25th, 2004 (Max ground sample distance = 0.62 m) and March 18th, 2005 (0.63 m) (Fig. 2). The 2005 image was used to map (i) new landslides, which are not visible (because they have not yet occurred) in the 2004 image, or (ii) reactivations of landslides already visible in 2004. Hence, only landslides that occurred or were reactivated between October 25th, 2004 and March 18th, 2005 were included in the inventories. Rainfall data from this

period is shown in Fig. 3. During this period, with the exception of October 2004 and February 2005, the monthly precipitation was higher than the 1970–2005 average.

The landslides were remotely recognized by identifying their typical diagnostic features such as scarps, concavo-convex profiles, irregular morphology, and cracks, in addition to land cover and drainage network modifications (Fig. 5). Moreover, damages to anthropogenic features such as roads (Fig. 5b), parcel borders and walls were also used to identify mass movements. However, most of the mapped landslides are rather shallow, and as observed from more recent Google Earth images, they are promptly leveled by farmers (Fig. 6) and/or smoothed by erosion. For most of the identified slope failures, the source areas could not be clearly distinguished from the transport and accumulation zones, and thus the entire landslide areas were mapped as unique polygons. Some of these polygons included two or more individual flows, but they were not distinguishable.

In total, we mapped 667 landslide areas (Fig. 7), 426 in AREA1 and 241 in AREA2. Most of them were classified as earth-flows. Landslides in AREA1 extend for 0.623 km^2 whereas those in AREA2 cover 0.436 km^2. The landslide density of AREA2 (4.2%) is more than double that of AREA1 (1.6%). Table 1 shows landslide areas according to lithology categories and investigated areas. The two dominant lithological units in AREA1 have very different susceptibilities to landsliding: clays host most landslides whereas marls and sandstones host much fewer. Such a marked influence of lithology is not recognizable in AREA2, where the frequency of the landslide area mostly agrees with that of the lithological category.

Landslide areas were mapped as vectors and then converted into a 5 m resolution grid layer, where the values 1 (one) and 0 (zero) were assigned to cells inside and outside landslides, respectively. The same resolution was also used for the predictor variables and the analysis of landslide susceptibility.

2.3. Predictor variables

Predictors of landslide occurrence are usually selected according to two main criteria: (i) their expected relationships with slope stability, and (ii) the quality and resolution of available data. Environmental attributes related to geology, topography/hydrology, soil properties and land use, are frequently employed for landslide susceptibility zonation (Costanzo et al., 2014; Conoscenti et al., 2015). For our study areas, detailed geological maps (scale 1:50,000) were available (Catalano et al., 2010; Di Stefano et al., 2013), as well as a high-resolution (2 m) LiDAR-derived DEM with a vertical error of 0.2 m (Regione Siciliana, 2010). On the other hand, available information on land cover and soil types had low spatial resolutions. Therefore, for this experiment, only lithology from the geological maps and 13 attributes derived from the DEM were employed as explanatory variables (Tables 1 and 2). These variables are often selected for the statistical assessment of landslide susceptibility (e.g., Rotigliano et al., 2011; Vorpahl et al., 2012; Felicísimo et al., 2013; Heckmann et al., 2014; Lombardo et al., 2015). Each variable is represented using a 5 m resolution raster grid layer. Lithology (LTL) has 15 categories; AREA1 has all of them while AREA2 has only nine of them (Table 1 and Fig. 4). The 13 terrain attributes were extracted from a 5 m resolution DEM, using the open source software SAGA-GIS (Olaya, 2004; Cimmery, 2010). This DEM was prepared by resampling the 2 m resolution DEM using bilinear interpolation. The 5 m resolution was chosen in order to compromise between spatial accuracy and reasonable computational times.

Elevation (*ELE*) simply corresponds to the values of the 5 m DEM. Altitude is often employed for landslide susceptibility zonation because of its correlation with climate (e.g. rainfall and temperature) and vegetation. Slope gradient (*SLO*) and aspect were calculated using the method of Zevenbergen and Thorne (1987). The catchment slope angle (*SLO_CAT*) was derived from *SLO*. Since slope aspect is a circular variable, it was divided into "northness" (*N*) and "eastness" (*E*), by using cosine and sine transformations, respectively; catchment northness (*N_CAT*) and eastness (*E_CAT*) were also calculated. To represent slope concavity/convexity, we included the convergence index (Köthe et al., 1996) and the topographic position index (*TPI*, Guisan et al.,

1999). The former was calculated at two different scales, by using a search radius of 1 cell (*CI*) or 10 cells (*CI*_10). *TPI* was computed using a radius of 100 m. The terrain ruggedness index, a measure of topographic heterogeneity (Riley et al., 1999), was also calculated at two distinct spatial scales: *TRI* using a radius of 1 cell, and *TRI*_10, using a radius of 10 cells. In order to account for potential soil saturation, the topographic wetness index (*TWI*; Beven and Kirkby, 1979) was also included. It is defined as

$$TWI = \ln(A/\tan\alpha) \qquad (1)$$

where A is the contributing area and α is the slope angle.

The variance inflation factor (*VIF*) was used to detect a potential strong correlation between two or more of the predictors. *VIF* was calculated for each of the terrain variables using the "usdm" package (Naimi, 2015), implemented in R software (R Core Team, 2015). As a *VIF* value greater than 10 is a signal of a serious collinearity problem (Heckmann et al., 2014; Jebur et al., 2014; Bui et al., 2015), this value was selected as the threshold to exclude collinear variables from the models.

2.4. Statistical modeling

In the first stage of the statistical analysis, we prepared a data matrix showing the values of the response and the predictor variables for each row. MARS was employed to model the statistical relationships between predictors and the presence/absence of landslides. MARS is a relatively new modeling technique that has been applied to various fields of environmental science (e.g. Leathwick et al., 2005; Naimi et al., 2011) and geomorphology (e.g. Gómez-Gutiérrez et al., 2009a,b, 2015; Shruthi et al., 2011). However, so far, only a few studies have used MARS for landslide susceptibility modeling (i.e. Vorpahl et al., 2012; Felicísimo et al., 2013; Conoscenti et al., 2015).

MARS is able to fit complex, non-linear relationships between response and explanatory variables while providing an interpretable model (Briand et al., 2004; Leathwick et al., 2005). This modeling

technique subdivides the range of the predictors into regions and fits a linear regression equation to each of them. The intervals are usually called "basis functions" (BFs), whereas the break values are called "knots". The general structure of MARS can be written as follows:

$$y = \alpha + \sum_{n=1}^{N} \beta_n h_n(x) \quad (2)$$

where y is the dependent variable, α is a constant, N is the number of terms, each formed by a coefficient β_n and $h_n(x)$ is an n-th single BF or a product of two or more basis functions of the independent variable x. An individual BF has the form max $(0, x - k)$ or max $(0, k - x)$, where x is a predictor and k is a knot. The MARS algorithm works in two stages: (i) generating a very complex and overfitted model by adding all possible terms; and (ii) decreasing the complexity of the model by identifying the subset of those terms that give the lowest value of the Generalized Cross-Validation parameter (GCV; Craven and Wahba, 1979). Please refer to Friedman (1991) for further details about the MARS algorithm.

The MARS modeling was performed using the "earth" package (Milborrow et al., 2011) of R software. In order to reduce the complexity of the models, the maximum degree of interaction was set equal to 1, thus avoiding terms given by combinations of two or more BFs. The software semi-automatically determined the maximum number of terms entering the MARS models. The "evimp" function of "earth" was employed to estimate the importance of each of the selected continuous predictors. This was evaluated according to the number of model subsets that include the variable. The higher this number is, the more important the contribution of the variable is. A model subset is the best subset of terms for each model size generated during the pruning pass (the second stage of MARS). Only subsets equal to or smaller than the final model are considered to evaluate predictor importance (Milborrow, 2015).

2.5. Sampling and validation strategy

The calibration of a MARS model needs a data set made of both positive and negative cases. No agreement exists on the proper amount of negative items to be used in multivariate statistical analyses of landslide susceptibility. In this study, samples with the same number of positive (landslide presence) and negative (landslide absence) cells were prepared for model calibration and validation. This was carried out using the random sampling tools in QGIS (QGIS Development Team, 2015) and R.

This experiment used three different data sets, two extracted from AREA1 and one from AREA2. The data sets of AREA1 were named A1a and A1b, whereas that of AREA2 was named A2. To evaluate the robustness of the procedure to changes of the input data, 10 subsets of cells were taken from the three data sets and were used to calibrate and validate the MARS models. The 10 subsets of A1a and A1b were divided into independent training and test samples. For each of them, positives cases were obtained by randomly splitting the 426 landslide polygons of AREA1 into 213 calibration and 213 validation landforms. A1a and A1b differ in the way negatives were selected. Following a method that has already been tested by Conoscenti et al. (2015), A1a used cells intersected by circles randomly distributed over the stable portion of the study area. These have a diameter of 20 m, which is approximately the average width of the landslide source areas. A circle simulates the initiation zone of a landslide moving towards all possible directions. On the other hand, absences of A1b samples were randomly selected as individual cells. Calibration and validation samples of A1a and A1b shared the same positives but included different negative cases. The A2 data set is made up of 10 subsets of cases, each joining all the presences occurring in AREA2 and the same number of randomly selected absences.

The independent training and test samples of both A1a and A1b were used to calibrate and validate 20 landslide susceptibility models. To evaluate LTL, half of these models were prepared without this predictor. Moreover, 20 MARS models were fitted to the complete (training + test samples) subsets of A1a and A1b and were then validated outside AREA1, using the A2 data set. Since not

all the LTL categories of AREA1 are present in AREA2, this external validation was carried out omitting LTL from the predictor variables. To summarize, six groups each made of 10 MARS replicates were calibrated and validated using the A1a, A1b and A2 data sets.

The accuracy of the landslide susceptibility models was assessed by using receiver operating characteristics (ROC) analysis (Goodenough et al., 1974; Lasko et al., 2005) and by calculating the area under the ROC curve (*AUC*) (Hanley and McNeil, 1982). In recent years, ROC curves and *AUC* statistics have been frequently adopted for the same purpose (e.g., Von Ruette et al., 2011; Schicker and Moon, 2012; Van Den Eeckhaut et al., 2012; Felicísimo et al., 2013; Conoscenti et al., 2015) and thus they may make it possible to compare the performance of different models and methods. An ROC curve plots the true positive rate (sensitivity) against the false negative rate (1 – specificity), at any given cut-off value. The upper-right portion of the curve corresponds to the lower cut-off scores. The closer the curve is to the upper-left corner, the larger the *AUC* value is, and the more accurate the model is. When a perfect classification between positives and negatives is achieved, the ROC curve passes through the point (0, 1) (*AUC* = 1). Conversely, an ROC curve close to the diagonal trend is produced when the model score shows no correlation with the dependent variable (*AUC* = 0.5). In this study, *AUC* values were used to measure both apparent (i.e. the fit to the training samples) and real accuracy (i.e. the skill in predicting the validation samples). Following the classification proposed by Hosmer and Lemeshow (2000), 0.7, 0.8 and 0.9 *AUC* thresholds were adopted to classify acceptable, excellent and outstanding performance, respectively. The differences in model accuracy were tested for statistical significance by submitting the *AUC* values obtained from all the model replicates to the Wilcoxon signed-rank test.

3. Results

3.1. Collinearity analysis

The variance inflation factor (*VIF*) was calculated to detect collinearity problems among the continuous predictor variables. The *VIF* values (Table 3) indicated *SLO* and *TRI* as highly correlated and thus they should not be included together in the models. At the same time, by excluding *TRI*, all *VIF* values lowered below 10, indicating the absence of a strong correlation between the remaining independent variables. Therefore, *TRI* was the only predictor omitted from the landslide susceptibility analysis.

3.2. Calibration and validation for AREA1

The *AUC* values of the models calibrated and validated for AREA1 (Fig. 8, Table 4) reveal that both A1a and A1b models that incorporate LTL showed an outstanding fit to the training data (mean *AUC* > 0.9). In predicting the test samples, the A1a and A1b models exhibited outstanding and excellent (mean *AUC* > 0.8) performance, respectively. Moreover, both A1a and A1b models prepared without LTL showed lower accuracy, although the goodness-of-fit on the calibration data and the predictive skill on the validation data remain excellent (mean *AUC* > 0.8). The differences of performance were tested for statistical significance. A significant difference in the calibration value of *AUC* was found between A1a models prepared with and without LTL (p = 0.002). The fit to the training data also became significantly better for the A1b models when LTL was included (p = 0.005). The *AUC* values for validation also confirm that incorporating LTL provides significantly better A1a and A1b models (p = 0.000). Regarding both goodness-of-fit and prediction skill, the A1a models perform significantly better than A1b models (p = 0.002 and 0.005, respectively). The same significant difference is also observed for the calibration *AUC* values (p = 0.002) and validation *AUC* values (p = 0.004) of models that do not include LTL.

3.3. Validation for AREA2

Fig. 9 summarizes the calibration and validation *AUC* values of the landslide susceptibility models trained on the entire A1a and A1b subsets of AREA1 and validated using the 10 subsets extracted from AREA2. Descriptive statistics of these *AUC* values are reported in Table 5. Both A1a and A1b models showed an excellent fit to their training data (mean *AUC* > 0.8) and an acceptable performance in predicting the AREA2 data (mean *AUC* > 0.7). However, the Wilcoxon signed-rank test demonstrated a significantly better fit to the learning data of the A1a models (p = 0.002). Conversely, no significant difference was found between A1a and A1b models when comparing their ability to predict presences and absences occurring on AREA2 (p = 0.846).

3.4. Landslide susceptibility maps of AREA1 and AREA2

Landslide susceptibility maps were prepared for both AREA1 and AREA2. Since the A1a models showed the best accuracy in AREA1 and the same accuracy as those of the A1b models in AREA2, the former were employed to produce the landslide susceptibility maps of both of the study areas. This was achieved by averaging the scores of the MARS replicates trained on the 10 A1a subsets for each pixel. Two landslide susceptibility maps were produced for AREA1, one including and one not including LTL as the predictor variable (Fig. 10a,b). The landslide susceptibility map of AREA2 was derived from the A1a models prepared without LTL (Fig. 10c).

3.5. Variable importance

Table 6 reports the number of model subsets that included each continuous variable. This number was calculated for all of the MARS replicates that were calibrated for the 10 A1a and A1b subsets. The predictor variables are ranked according to the sum of model subsets including them. *TRI*_10 and *TPI* were the most frequently selected. Comparing the model subsets of the 10 replicates, a higher robustness of the A1b models arises. Indeed, both the size (i.e. number of terms) of the final

models and the number of times each variable enters the model subsets of the A1b replicates varies very little compared to the 10 A1a models.

4. Discussion

In this section, the results of our experiment are discussed and compared with the findings of other landslide susceptibility studies that exploited *AUC* statistics to measure model performance, but made different choices regarding one or more of the following: (i) type of landslides, (ii) type of mapping unit, (iii) spatial scale, (iv) mapping technique, (v) set of predictor variables, (vi) type of instability diagnostic features (e.g., landslide areas, scarps, and deposits), (vii) sampling strategy, (viii) technique employed to assess the model performance and (ix) model validation scheme (i.e. temporal and/or spatial partition of the calibration and validation subsets). The excellent to outstanding goodness-of-fit and predictive skill of the models observed in AREA1 demonstrate the validity of our approach and confirm that accurate predictions of landslide occurrence may be obtained from a MARS model. These findings may be explained by the ability of MARS to detect complex non-linear relationships such as those between terrain attributes and proneness to landslides. Compared to other statistical methods frequently employed with this aim, such as discriminant analysis (e.g., Carrara, 1983; Carrara et al., 1991, 2008; Guzzetti et al., 2006) and logistic regression (e.g., Ohlmacher et al., 2003; Lee, 2005; Nefeslioglu et al., 2008; Costanzo et al., 2014), MARS has the advantage of dividing the range of the predictor variables into regions and fitting each of them with a separate linear regression. However, in order to prevent overfitting to the calibration data, which would produce a very poor prediction of landslide distribution outside the training samples, the MARS model should be kept as simple as possible. Here, this was achieved by preparing models with terms involving only one variable.

The high accuracy of the models also demonstrates the reliability of our landslide inventory, which was prepared exclusively with a visual analysis of Google Earth images, without any field checking. This result suggests that effective landslide susceptibility maps may also be prepared for areas, such

as developing countries, where high-accuracy remote data are not readily available and field surveys may be difficult to perform (e.g. for security problems). Moreover, due to the availability of multi-temporal images, which are often more frequent than other remote data provided by national or regional institutions, landslide inventories and susceptibility maps may be updated in a short time and at essentially no cost. Among the studies exploiting landslide inventory that were established, at least in part, from Google Earth images, Van Den Eeckhaut et al. (2012) found a similar fit to the validation data (*AUC* = 0.854–0.923) whereas Schicker and Moon (2012) obtained a quite lower performance (*AUC* = 0.621–0.712). However, these studies employed different modeling techniques and sets of predictors and they were carried out at the regional scale. Moreover, the predictive skill of the models prepared by Schicker and Moon (2012) was measured outside the calibration area. A validation with data from other study areas may lead to significantly lower accuracy, in particular if training and test areas have different environmental characteristics. In this study, models trained for AREA1 achieved *AUC* values of 0.718 to 0.763 when validated in AREA2. This performance is similar to that of other studies attempting a "true" external validation, such as Von Ruette et al. (2011) and Pradhan et al. (2010) who obtained *AUC* values of 0.69–0.82 and 0.732–0.792, respectively. A notable exception is represented by Lombardo et al. (2014), who also achieved excellent performance outside the training area (Mean *AUC* = 0.822). However, these authors analyzed two adjacent small basins with similar geological and geomorphological conditions, whereas our two experimental areas are quite diverse. Nevertheless, most of the landslides mapped in AREA2 occur where predicted susceptibility is high (Fig. 10c). Furthermore, the overall stability of the predictive performance demonstrates that our procedure is also robust when validation is performed outside the calibration area (Table 5).

In our experiment, LTL demonstrated substantially improving the predictive ability of the landslide susceptibility models, confirming the observation of Felicísimo et al. (2013) in Spain. This result was expected but not obvious because identifying a robust relationship between LTL and landslide susceptibility may depend on various factors, including the accuracy of the available lithological

map, the criteria used for its classification and the quality of the landslide inventory. However, due to the different geological settings of the study areas, we could not incorporate LTL in the models tested in AREA2. Attempts to include this variable by identifying groups of LTL categories occurring in both areas led to worse performance than that achievable without using LTL.

The models incorporating LTL and prepared using random circles to sample the landslide absences (i.e. the A1a models) showed a fit to the learning data (*AUC* = 0.929–0.948) and to the validation data (*AUC* = 0.891–0.929) that is very similar to the result of Conoscenti et al. (2015). These authors employed the same method of negative selection from a basin close to our study areas, achieving calibration and validation *AUC* values of 0.853–0.912 and 0.918–0.941, respectively. In AREA1, the performance of the A1a models was significantly better than that of the A1b models. This is probably due to a stronger link between the A1a models and their calibration data, which are less uniformly distributed than those of the A1b data set. Corroborating this hypothesis, the performance difference between A1a and A1b models observed for calibration data is larger than that calculated for validation data (Table 4 and Fig. 8). Moreover, no significant difference in accuracy was found outside the calibration area (Table 5 and Fig. 9). A stronger relationship with the training data may also explain: (i) the higher variability of accuracy shown in AREA2 by the A1a replicates (Table 5 and Fig. 9); and (ii) their slightly lower robustness concerning the number of model subsets including each predictor and the final size of the models (Table 6). The validation results thus suggest that if the objective is to predict unknown landslides in the same area where calibration is performed then selecting the absences using random circles may yield a better accuracy of the models. On the other hand, if the aim is to evaluate landslide susceptibility outside the calibration area, both the methods of negative selection are suitable for predicting landslide occurrence even if a slightly higher robustness of the models may be expected if negatives are sampled using random points.

5. Concluding remarks

In this experiment, we assessed landslide susceptibility in two study areas in Sicily (Italy), by using MARS as a modeling technique. The models were based on inventories of rainfall-triggered landslides that occurred during winter 2004–2005, which were identified exclusively by visual interpretation of Google Earth images. We evaluated the importance of LTL by measuring the accuracy of models including or not including this variable. Calibration and validation data sets were prepared using two different methods of landslide absences sampling.

We found excellent to outstanding accuracy of the models when calibration and validation were performed with data from the same area (i.e. AREA1). Performance of the models was significantly higher when LTL was used as a predictor variable. Prediction skill decreased to an acceptable level when validation was performed outside the training area (i.e. in AREA2). Validation in AREA1 also showed that extraction of the negatives from circles with the same diameter as landslide source areas provided significantly better model accuracy than absences selected as individual grid cells. Conversely, no significant difference of predictive ability was observed when landslide susceptibility models were exported to AREA2, although a slightly higher robustness of the models was found when negatives were sampled as random cells.

The experiment demonstrated the consistence of our landslide inventories, suggesting that reliable landslide susceptibility models may be prepared for large or hardly accessible areas, at a low cost, if Google Earth images of proper resolution are available. The results showed that when seeking to predict unknown landslides in the same area where calibration is performed, then selecting the absences with random circles seems to yield a better predictive performance.

The adopted procedure allowed us to achieve a reliable assessment of landslide susceptibility in a short time by using Google Earth images, a lithological map and a set of DEM derivatives. The method is relatively simple and may be reproduced by employing open source software and environmental data, which are often available for free. This approach may help land management

agencies to achieve an accurate prediction of landslide occurrence and establish preventive and mitigation measures.

Acknowledgments

This research was developed in the framework of the projects: FLUMEN (project number: 318969), funded by the EU (call identifier: FP7-PEOPLE-2012-IRSES); SUFRA_SICILIA funded by the Department of Earth and Marine Sciences of the University of Palermo and by the Assessorato Regionale Territorio e Ambiente della Regione Sicilia. Valerio Agnesi has participated in the final discussion of the data; the other authors have collaborated together during all of the research phases. The authors wish to thank Dr. Marilena Ciaccio, who participated in the landslide mapping of the first study area. The authors also thank the editor Prof. Takashi Oguchi and the two anonymous reviewers for their suggestions and comments. Finally, the authors wish to thank Cassandra Funsten for her revision of the paper's English.

References

Agnesi, V., Macaluso, T., Monteleone, S., Pipitone, G., 1982. Indagine geomorfologica ed analisi statistica dei dissesti dell'alto bacino del Fiume San Leonardo (Sicilia occidentale). Geol. Appl. e Idrogeol. 17(1), 243–271

Aleotti, P., Chowdhury, R., 1999. Landslide hazard assessment: summary review and new perspectives. Bull. Eng. Geol. Environ. 58, 21–44. doi:10.1007/s100640050066

Atkinson, P.M., Massari, R., 2011. Autologistic modelling of susceptibility to landsliding in the Central Apennines, Italy. Geomorphology 130, 55–64. doi:10.1016/j.geomorph.2011.02.001

Beven, K.J., Kirkby, M.J., 1979. A physically based variable contributing area model of basin hydrology. Hydrol. Sci. Bull. 24, 43–69.

Borrelli, L., Cofone, G., Coscarelli, R., Gullà, G., 2014. Shallow landslides triggered by consecutive rainfall events at Catanzaro strait (Calabria–Southern Italy). J. Maps 37–41. doi:10.1080/17445647.2014.943814

Brabb, E.E., 1984. Innovative approaches to landslide hazard and risk mapping, in: Proceedings 4th International Symposium on Landslides Vol. 1. Toronto, Canada, pp. 307–324.

Brenning, A., 2005. Spatial prediction models for landslide hazards: review, comparison and evaluation. Nat. Hazards Earth Syst. Sci. 5, 853–862. doi:10.5194/nhess-5-853-2005

Briand, L.C., Freimut, B., Vollei, F., 2004. Using multiple adaptive regression splines to support decision making in code inspections. J. Syst. Softw. 73, 205–217. doi:10.1016/j.jss.2004.01.015

Bui, D.T., Tuan, T.A., Klempe, H., Pradhan, B., Revhaug, I., 2015. Spatial prediction models for shallow landslide hazards : a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. Landslides. doi:10.1007/s10346-015-0557-6

Cama, M., Lombardo, L., Conoscenti, C., Agnesi, V., Rotigliano, E., 2015. Predicting storm-triggered debris flow events: application to the 2009 Ionian Peloritan disaster (Sicily, Italy). Nat. Hazards Earth Syst. Sci. 15, 1785–1806. doi:10.5194/nhess-15-1785-2015

Carrara, A., 1983. Multivariate models for landslide hazard evaluation. J. Int. Assoc. Math. Geol. 15, 403–426.

Carrara, A., Cardinali, M., Detti, R., Guzzetti, F., Pasqui, V., Reichenbach, P., 1991. GIS techniques and statistical models in evaluating landslide hazard. Earth Surf. Process. Landforms 16, 427–445.

Carrara, A., Cardinali, M., Guzzetti, F., Reichenbach, P., 1995. GIS technology in mapping landslide hazard, in: Carrara, A., Guzzetti, F. (Eds.), Geographical Information Systems in Assessing Natural Hazards. Kluwer, Dordrecht, pp. 135–175.

Carrara, A., Crosta, G., Frattini, P., 2008. Comparing models of debris-flow susceptibility in the alpine environment. Geomorphology 94, 353–378.

Catalano, R., Avellone, G., Basilone, L., Sulli, A., 2010. Carta Geologica d'Italia alla scala 1:50.000, Foglio 607 - Corleone [WWW Document]. ISPRA – Serv. Geol. D'Italia. URL http://www.isprambiente.gov.it/Media/carg/607_CORLEONE/Foglio.html (accessed 1.14.14).

Chung, C.-J.F., Fabbri, A.G., 2003. Validation of Spatial Prediction Models for Landslide Hazard Mapping. Nat. Hazards 30, 451–472. doi:10.1023/B:NHAZ.0000007172.62651.2b

Cimmery, V., 2010. SAGA User Guide, updated for SAGA version 2.0.5.

Clerici, A., Perego, S., Tellini, C., Vescovi, P., 2002. A procedure for landslide susceptibility zonation by the conditional analysis method. Geomorphology 48, 349–364. doi:10.1016/S0169-555X(02)00079-X

Conforti, M., Pascale, S., Robustelli, G., Sdao, F., 2014. Evaluation of prediction capability of the artificial neural networks for mapping landslide susceptibility in the Turbolo River catchment (northern Calabria, Italy). Catena 113, 236–250. doi:10.1016/j.catena.2013.08.006

Conoscenti, C., Ciaccio, M., Caraballo-Arias, N.A., Gómez-Gutiérrez, Á., Rotigliano, E., Agnesi, V., 2015. Assessment of susceptibility to earth-flow landslide using logistic regression and multivariate adaptive regression splines: A case of the Belice River basin (western Sicily, Italy). Geomorphology 242, 49–64. doi:10.1016/j.geomorph.2014.09.020

Costanzo, D., Cappadonia, C., Conoscenti, C., Rotigliano, E., 2012a. Exporting a Google EarthTM aided earth-flow susceptibility model: a test in central Sicily. Nat. Hazards 61, 103–114. doi:10.1007/s11069-011-9870-0

Costanzo, D., Rotigliano, E., Irigaray, C., Jiménez-Perálvarez, J.D., Chacón, J., 2012b. Factors selection in landslide susceptibility modelling on large scale following the GIS matrix method: application to the river Beiro basin (Spain). Nat. Hazards Earth Syst. Sci. 12, 327–340. doi:10.5194/nhess-12-327-2012

Costanzo, D., Chacón, J., Conoscenti, C., Irigaray, C., Rotigliano, E., 2014. Forward logistic regression for earth-flow landslide susceptibility assessment in the Platani river basin (southern Sicily, Italy). Landslides 11, 639–653. doi:10.1007/s10346-013-0415-3Craven, P., Wahba, G., 1979. Smoothing noisy data with spline functions. Numer. Math. 31, 377–403.

Di Stefano, P., Renda, P., Zarcone, G., Nigro, F., Cacciatore, M.S., 2013. Carta Geologica d'Italia alla scala 1:50.000, Foglio 619 - S.Margherita di Belice [WWW Document]. ISPRA – Serv. Geol. D'Italia. URL http://www.isprambiente.gov.it/Media/carg/619_SMARGHERITA_BELICE/Foglio.html (accessed 6.22.15).

European Environment Agency, 2010. Corine Land Cover 2006 [WWW Document]. URL http://www.eea.europa.eu/data-and-maps/data/corine-land-cover-2006-raster (accessed 1.1.15).

Felicísimo, Á.M., Cuartero, A., Remondo, J., Quirós, E., 2013. Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: a comparative study. Landslides 10, 175–189. doi:10.1007/s10346-012-0320-1

Friedman, J.H., 1991. Multivariate adaptive regression splines. Ann. Stat. 19, 1–141.

Goetz, J.N., Brenning, A., Petschko, H., Leopold, P., 2015. Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. Comput. Geosci. 81, 1–11. doi:10.1016/j.cageo.2015.04.007

Gómez-Gutiérrez, Á., Schnabel, S., Felicísimo, Á.M., 2009a. Modelling the occurrence of gullies in rangelands of southwest Spain. Earth Surf. Process. Landforms 34, 1894–1902. doi:10.1002/esp1881

Gómez-Gutiérrez, Á., Schnabel, S., Lavado Contador, F., 2009b. Using and comparing two nonparametric methods (CART and MARS) to model the potential distribution of gullies. Ecol. Modell. 220, 3630–3637. doi:10.1016/j.ecolmodel.2009.06.020

Gómez-Gutiérrez, Á., Conoscenti, C., Angileri, S.E., Rotigliano, E., Schnabel, S., 2015. Using topographical attributes to evaluate gully erosion proneness (susceptibility) in two mediterranean basins: advantages and limitations. Nat. Hazards. doi:10.1007/s11069-015-1703-0

Goodenough, D.J., Rossmann, K., Lusted, L.B., 1974. Radiographic applications of receiver operating characteristic (ROC) curves. Radiology 110, 89–95.

Guisan, A., Weiss, S.B., Weiss, A.D., 1999. GLM versus CCA spatial modeling of plant species distribution. Plant Ecol. 143, 107–122.

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. doi:10.1016/S0169-555X(99)00078-1

Guzzetti, F., Reichenbach, P., Ardizzone, F., Cardinali, M., Galli, M., 2006. Estimating the quality of landslide susceptibility models. Geomorphology 81, 166–184.

Guzzetti, F., Mondini, A.C., Cardinali, M., Fiorucci, F., Santangelo, M., Chang, K.-T., 2012. Landslide inventory maps: New tools for an old problem. Earth-Science Rev. 112, 42–66. doi:10.1016/j.earscirev.2012.02.001

Hanley, J.A., McNeil, B.J., 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143, 29–36.

Heckmann, T., Gegg, K., Gegg, a., Becht, M., 2014. Sample size matters: investigating the effect of sample size on a logistic regression susceptibility model for debris flows. Nat. Hazards Earth Syst. Sci. 14, 259–278. doi:10.5194/nhess-14-259-2014

Hosmer, D.W., Lemeshow, S., 2000. Applied logistic regression, Wiley Series in Probability and Statistics, Wiley series in probability and statistics: Texts and references section. Wiley. doi:10.1198/tech.2002.s650

Jebur, M.N., Pradhan, B., Tehrany, M.S., 2014. Optimization of landslide conditioning factors using very high-resolution airborne laser scanning (LiDAR) data at catchment scale. Remote Sens. Environ. 152, 150–165. doi:10.1016/j.rse.2014.05.013

Köthe, R., Gehrt, E., Böhner, J., 1996. Automatische Reliefanalyse für geowissenschaftliche Anwendungen— derzeitiger Stand undWeiterentwicklungen des Programms SARA. Arbeitshefte Geol. 1, 31–37.

Lasko, T.A., Bhagwat, J.G., Zou, K.H., Ohno-Machado, L., 2005. The use of receiver operating characteristic curves in biomedical informatics. J. Biomed. Inform. 38, 404–415. doi:10.1016/j.jbi.2005.02.008

Leathwick, J.R., Rowe, D., Richardson, J., Elith, J., Hastie, T., 2005. Using multivariate adaptive regression splines to predict the distributions of New Zealand's freshwater diadromous fish. Freshw. Biol. 50, 2034–2052. doi:10.1111/j.1365-2427.2005.01448.x

Lee, S., 2005. Application and cross-validation of spatial logistic multiple regression for landslide susceptibility analysis. Geosci. J. 9, 63–71. doi:10.1007/BF02910555

Lombardo, L., Cama, M., Maerker, M., Rotigliano, E., 2014. A test of transferability for landslides susceptibility models under extreme climatic events: application to the Messina 2009 disaster. Nat. Hazards 1951–1989. doi:10.1007/s11069-014-1285-2

Lombardo, L., Cama, M., Conoscenti, C., Märker, M., Rotigliano, E., 2015. Binary logistic regression versus stochastic gradient boosted decision trees in assessing landslide susceptibility for multiple-occurring landslide events: application to the 2009 storm event in Messina (Sicily, southern Italy). Nat. Hazards. doi:10.1007/s11069-015-1915-3

Milborrow, S., 2015. Notes on the earth package [WWW Document]. URL http://www.milbo.org/doc/earth-notes.pdf (accessed 7.15.15).

Milborrow, S., Hastie, T., Tibshirani, R., 2011. Earth: Multivariate Adaptive Regression Spline Models. R Software Package.

Naimi, B., 2015. Uncertainty analysis for species distribution models. R Software Package.

Naimi, B., Skidmore, A.K., Groen, T. a., Hamm, N. a. S., 2011. Spatial autocorrelation in predictors reduces the impact of positional uncertainty in occurrence data on species distribution modelling. J. Biogeogr. 38, 1497–1509. doi:10.1111/j.1365-2699.2011.02523.x

Nefeslioglu, H.A., Gokceoglu, C., Sonmez, H., 2008. An assessment on the use of logistic regression and artificial neural networks with different sampling strategies for the preparation of landslide susceptibility maps. Eng. Geol. 97, 171–191. doi:10.1016/j.enggeo.2008.01.004

Ohlmacher, G.C., Davis, J.C., 2003. Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA. Eng. Geol. 69, 331–343. doi:10.1016/S0013-7952(03)00069-3

Olaya, V., 2004. A gentle introduction to SAGA GIS. Göettingen, Germany.

Pradhan, B., Lee, S., Buchroithner, M.F., 2010. A GIS-based back-propagation neural network model and its cross-application and validation for landslide susceptibility analyses. Comput. Environ. Urban Syst. 34, 216–235. doi:10.1016/j.compenvurbsys.2009.12.004

QGIS Development Team, 2015. QGIS User Guide - Release 2.8.

R Core Team, 2015. An Introduction to R. URL https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf (accessed 7.15.2015).

Regione Siciliana, 2010. Modello digitale del terreno (MDT) 2m x 2m Regione Siciliana - ATA 2007-2008 [WWW Document]. URL http://www.sitr.regione.sicilia.it

Riley, S.J., DeGloria, S.D., Elliot, R., 1999. A terrain ruggedness index that quantifies topographic heterogeneity. Intermt. J. Sci. 5, 23–27.

Rotigliano, E., Agnesi, V., Cappadonia, C., Conoscenti, C., 2011. The role of the diagnostic areas in the assessment of landslide susceptibility models: a test in the sicilian chain. Nat. Hazards 58, 981–999. doi:10.1007/s11069-010-9708-1

Rotigliano, E., Cappadonia, C., Conoscenti, C., Costanzo, D., Agnesi, V., 2012. Slope units-based flow susceptibility model: using validation tests to select controlling factors. Nat. Hazards 61, 143–153. doi:10.1007/s11069-011-9846-0

Schicker, R., Moon, V., 2012. Comparison of bivariate and multivariate statistical approaches in landslide susceptibility mapping at a regional scale. Geomorphology 161-162, 40–57. doi:10.1016/j.geomorph.2012.03.036

Shruthi, R.B.V., Kerle, N., Jetten, V., 2011. Object-based gully feature extraction using high spatial resolution imagery. Geomorphology 134, 260–268. doi:10.1016/j.geomorph.2011.07.003

Van Den Eeckhaut, M., Vanwalleghem, T., Poesen, J., Govers, G., Verstraeten, G., Vandekerckhove, L., 2006. Prediction of landslide susceptibility using rare events logistic regression: A case-study in the Flemish Ardennes (Belgium). Geomorphology 76, 392–410. doi:10.1016/j.geomorph.2005.12.003

Van Den Eeckhaut, M., Hervás, J., Jaedicke, C., Malet, J.P., Montanarella, L., Nadim, F., 2012. Statistical modelling of Europe-wide landslide susceptibility using limited landslide inventory data. Landslides 9, 357–369. doi:10.1007/s10346-011-0299-z

Van Westen, C.J., Asch, T.W.J., Soeters, R., 2005. Landslide hazard and risk zonation—why is it still so difficult? Bull. Eng. Geol. Environ. 65, 167–184. doi:10.1007/s10064-005-0023-0

Van Westen, C.J., Castellanos, E., Kuriakose, S.L., 2008. Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. Eng. Geol. 102, 112–131. doi:10.1016/j.enggeo.2008.03.010

Vergari, F., Della Seta, M., Del Monte, M., Fredi, P., Lupia Palmieri, E., 2011. Landslide susceptibility assessment in the Upper Orcia Valley (Southern Tuscany, Italy) through conditional analysis: a contribution to the unbiased selection of causal factors. Nat. Hazards Earth Syst. Sci. 11, 1475–1497. doi:10.5194/nhess-11-1475-2011

Von Ruette, J., Papritz, A., Lehmann, P., Rickli, C., Or, D., 2011. Spatial statistical modeling of shallow landslides-Validating predictions for different landslide inventories and rainfall events. Geomorphology 133, 11–22. doi:10.1016/j.geomorph.2011.06.010

Vorpahl, P., Elsenbeer, H., Märker, M., Schröder, B., 2012. How can statistical models help to determine driving factors of landslides? Ecol. Modell. 239, 27–39. doi:10.1016/j.ecolmodel.2011.12.007

Zevenbergen, L.W., Thorne, C.R., 1987. Quantitative analysis of land surface topography. Earth Surf. Process. Landforms 12, 47–56.

Zhang, J., Gurung, D.R., Liu, R., Murthy, M.S.R., Su, F., 2015. Abe Barek landslide and landslide susceptibility assessment in Badakhshan Province, Afghanistan. Landslides 12, 597–609. doi:10.1007/s10346-015-0558-5

Captions

Fig. 1. Location of the study areas.

Fig. 2. Elevation and shaded relief map of the study areas.

Fig. 3. Box plot summarizing the average monthly rainfall recorded in Corleone (588 m a.s.l.). Grey dots highlight rainfall related to the triggering of the mapped landslides.

Fig. 4. LTL categories of AREA1 (a) and AREA2 (b).

Fig. 5. Google Earth views of some of the landslides mapped in AREA1 (a–c) and AREA2 (d). Locations of the images are shown in Fig. 2.

Fig. 6. Google Earth views of a portion of AREA1 dated October 25th 2004 (a), March 18th 2005 (b) and September 30th 2006 (c), showing some of the mapped landslides leveled by farmers after March 18th 2005. The location of the images is shown in Fig. 2.

Fig. 7. Landslide maps of AREA1 (a) and AREA2 (b).

Fig. 8. Box plot summarizing *AUC* values of MARS replicates calibrated and validated in AREA1.

Fig. 9. Box plot summarizing *AUC* values of MARS replicates calibrated for AREA1 and validated in AREA2.

Fig. 10. Landslide susceptibility maps. (a) AREA1 from A1a models including LTL as a predictor variable; (b) not including LTL, and (c) AREA2 from A1a models not including LTL.

Table 1. Relative frequency of LTL categories and landslides in AREA1 and AREA2.

Table 2. Continuous predictor variables and their statistics for AREA1 and AREA2.

Table 3. Values of the variance inflation factor (*VIF*) calculated including (*VIF*1) and not including (*VIF*2) the variable *TRI*.

Table 4. Descriptive statistics of *AUC* values computed from MARS models calibrated and validated using the A1a and A1b training and test samples.

Table 5. Descriptive statistics of *AUC* values computed from MARS models calibrated and validated using the A1a, A1b and A2 subsets.

Table 6. Number of model subsets including each predictor, calculated for all MARS replicates calibrated for the A1a and A1b subsets.

Table 1.

Lithology (LTL) classes		AREA1 Relative frequency		AREA2 Relative frequency	
Code	Description	LTL classes	Landslide area	LTL classes	Landslide area
LTL01	Alluvial deposits	18.5%	9.3%	4.6%	2.4%
LTL02	Eluvial-colluvial deposits	7.1%	6.4%	1.6%	4.9%
LTL03	Gravels and sands	1.3%	0.0%	0.0%	0.0%
LTL04	Limestones	0.5%	0.1%	1.2%	0.2%
LTL05	Marls	0.1%	0.3%	43.3%	48.6%
LTL06	Marls and marly limestones	2.8%	2.0%	0.0%	0.0%
LTL07	Sandstones and sandy clays	5.1%	4.2%	0.0%	0.0%
LTL08	Selenitic gypsum	0.6%	0.1%	0.0%	0.0%
LTL09	Talus deposits	0.3%	0.0%	0.4%	0.0%
LTL10	Ancient landslides	4.5%	14.6%	28.0%	30.5%
LTL11	Calcarenites	0.1%	0.0%	15.4%	10.9%
LTL12	Clays	31.4%	53.3%	1.4%	1.2%
LTL13	Marls and sandstones	25.9%	3.4%	0.0%	0.0%
LTL14	Clays and sandy clays	1.2%	5.9%	4.2%	1.4%
LTL15	Conglomerates	0.5%	0.4%	0.0%	0.0%

Table 2.

Continuous predictor variables		**Descriptive statistics over AREA1**				**Descriptive statistics over AREA2**			
Code	Description	Mean	St.dev.	Min	Max	Mean	St.dev.	Min	Max
ELE	Elevation [m asl]	371.636	68.876	217.662	559.480	481.228	78.848	317.128	713.598
SLO	Slope angle [°]	9.507	6.010	0.000	69.231	10.820	5.608	0.013	54.071
SLO_CAT	Catchment slope angle [°]	9.621	4.784	0.011	51.333	11.867	4.981	0.029	48.170
N	Northness	0.141	0.706	-1.000	1.000	0.205	0.756	-1.000	1.000
E	Eastness	0.004	0.708	-1.000	1.000	-0.158	0.602	-1.000	1.000
N_CAT	Catchment northness	0.167	0.698	-1.000	1.000	0.245	0.785	-1.000	1.000
E_CAT	Catchment eastness	0.237	0.655	-1.000	1.000	0.173	0.542	-1.000	1.000
CI	Convergence index [°]	0.006	10.540	-95.114	97.412	-0.040	8.828	-93.954	93.297
*CI*_10	Convergence index with search radius of 50 m [°]	0.045	17.825	-66.277	89.583	-0.400	16.064	-58.489	76.778
TPI	Topographic position index [m]	-0.001	0.996	-8.175	9.610	-0.035	0.795	-6.354	7.097
TRI	Terrain ruggedness index [m]	0.614	0.407	0.003	15.285	0.698	0.390	0.002	4.902
*TRI*_10	Terrain ruggedness index with search radius of 50 m [m]	4.216	2.391	0.122	24.635	4.737	2.111	0.373	21.292
TWI	Topographic wetness index [m]	8.477	1.921	3.168	24.493	8.371	1.665	3.585	19.047

Table 3.

Variables	*VIF*1	*VIF*2
ELE	1.10	1.09
SLO	113.07	6.64
SLO_CAT	4.10	3.88
N	9.01	8.26
E	1.00	1.00
N_CAT	9.19	8.41
E_CAT	1.08	1.09
CI	1.25	1.30
*CI*_10	3.67	3.69
TPI	3.55	3.65
TRI	100.77	----
*TRI*_10	7.04	6.64
TWI	2.68	2.63

Table 4.

Group of replicates	Lithology	Calibration	Validation	*AUC* values of 10 MARS replicates			
				Mean	St.dev.	Min	Max
(i)	YES	A1a training samples	A1a training samples	0.937	0.005	0.929	0.948
		A1a training samples	A1a test samples	0.909	0.011	0.891	0.929
(ii)		A1b training samples	A1b training samples	0.911	0.008	0.902	0.921
		A1b training samples	A1b test samples	0.889	0.009	0.877	0.903
(iii)	NO	A1a training samples	A1a training samples	0.877	0.012	0.860	0.895
		A1a training samples	A1a test samples	0.852	0.015	0.829	0.875
(iv)		A1b training samples	A1b training samples	0.861	0.010	0.843	0.875
		A1b training samples	A1b test samples	0.839	0.012	0.822	0.863

Table 5.

Group of replicates	Calibration	Validation	*AUC* values of 10 MARS replicates			
			Mean	St.dev.	Min	Max
(v)	A1a subsets	A1a subsets	0.869	0.004	0.864	0.876
	A1a subsets	A2 subsets	0.744	0.014	0.718	0.763
(vi)	A1b subsets	A1b subsets	0.855	0.002	0.853	0.858
	A1b subsets	A2 subsets	0.744	0.006	0.736	0.752

Table 6.

Number of model subsets that include the variable													
	MARS replicates calibrated for the A1a subsets										Descriptive statistics		
Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Sum**	**Mean**	**St.dev.**
*TRI*_10	21	18	18	19	20	18	20	19	19	19	191	19.1	1.0
TPI	19	16	18	17	18	16	20	18	17	19	178	17.8	1.3
ELE	17	15	14	15	17	15	17	16	16	15	157	15.7	1.1
N_CAT	18	13	15	16	16	14	16	14	14	17	153	15.3	1.6
SLO_CAT	16	12	16	7	15	13	18	17	13	16	143	14.3	3.2
TWI	13	11	12	13	12	10	13	7	11	11	113	11.3	1.8
E_CAT	15	9	9	8	11	7	15	11	8	8	101	10.1	2.9
*CI*_10	5	0	0	11	10	0	10	0	15	13	64	6.4	6.1
CI	6	6	7	0	7	6	8	13	4	7	64	6.4	3.2
SLO	0	8	0	0	5	0	0	0	0	0	13	1.3	2.8
N	0	0	0	0	0	0	0	0	0	0	0	0	0.0
E	0	0	0	0	0	0	0	0	0	0	0	0	0.0
Size of the final model	**21**	**18**	**19**	**19**	**20**	**18**	**21**	**19**	**19**	**20**	**194**	**19.4**	**1.1**

	MARS replicates calibrated for the A1b subsets										Descriptive statistics		
Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Sum**	**Mean**	**St.dev.**
*TRI*_10	18	18	19	18	19	19	18	18	18	18	183	18.3	0.5
TPI	16	16	17	16	17	17	16	16	16	16	163	16.3	0.5
N_CAT	14	14	16	15	13	15	15	15	15	14	146	14.6	0.8
ELE	13	13	15	14	16	14	14	14	14	13	140	14	0.9
TWI	11	12	14	13	15	13	13	13	13	12	129	12.9	1.1
SLO_CAT	15	15	11	10	11	16	10	10	10	15	123	12.3	2.6
CI	6	6	7	6	7	7	6	6	6	6	63	6.3	0.5
E_CAT	5	5	6	5	6	6	5	5	5	5	53	5.3	0.5
*CI*_10	0	0	0	0	0	0	0	0	0	0	0	0	0.0
E	0	0	0	0	0	0	0	0	0	0	0	0	0.0
N	0	0	0	0	0	0	0	0	0	0	0	0	0.0
SLO	0	0	0	0	0	0	0	0	0	0	0	0	0.0
Size of the final model	**18**	**18**	**19**	**18**	**19**	**19**	**18**	**18**	**18**	**18**	**183**	**18.3**	**0.5**

Figure 1

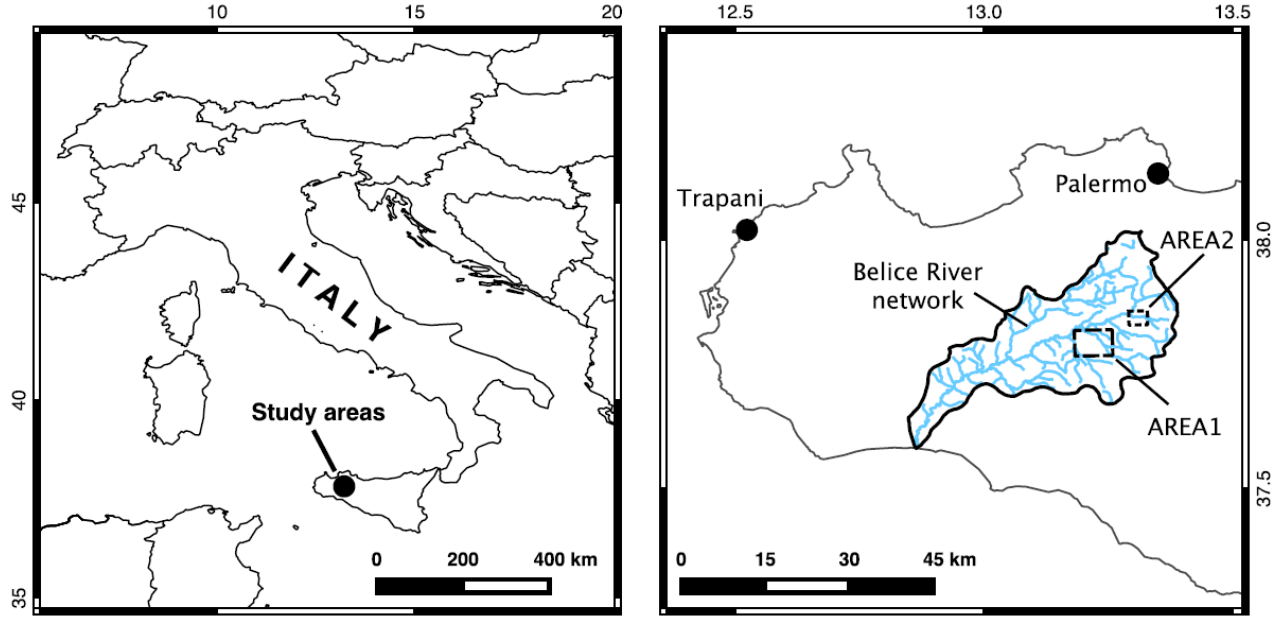


Figure 2

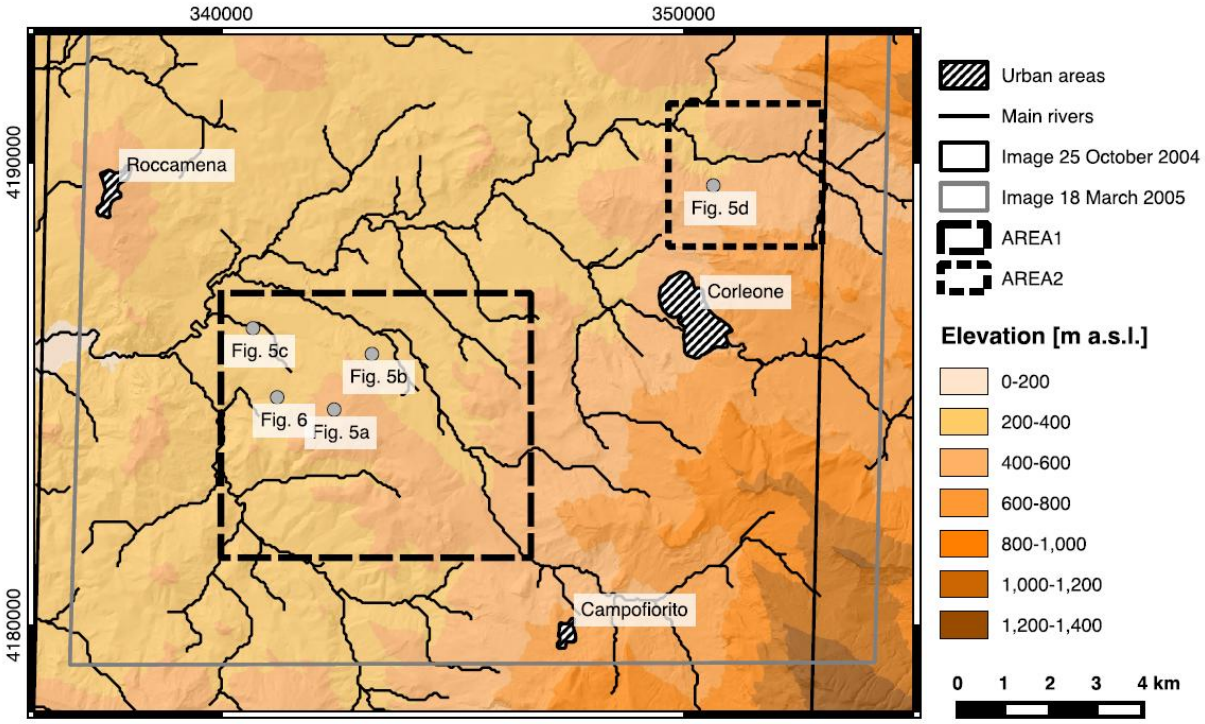


Figure 3

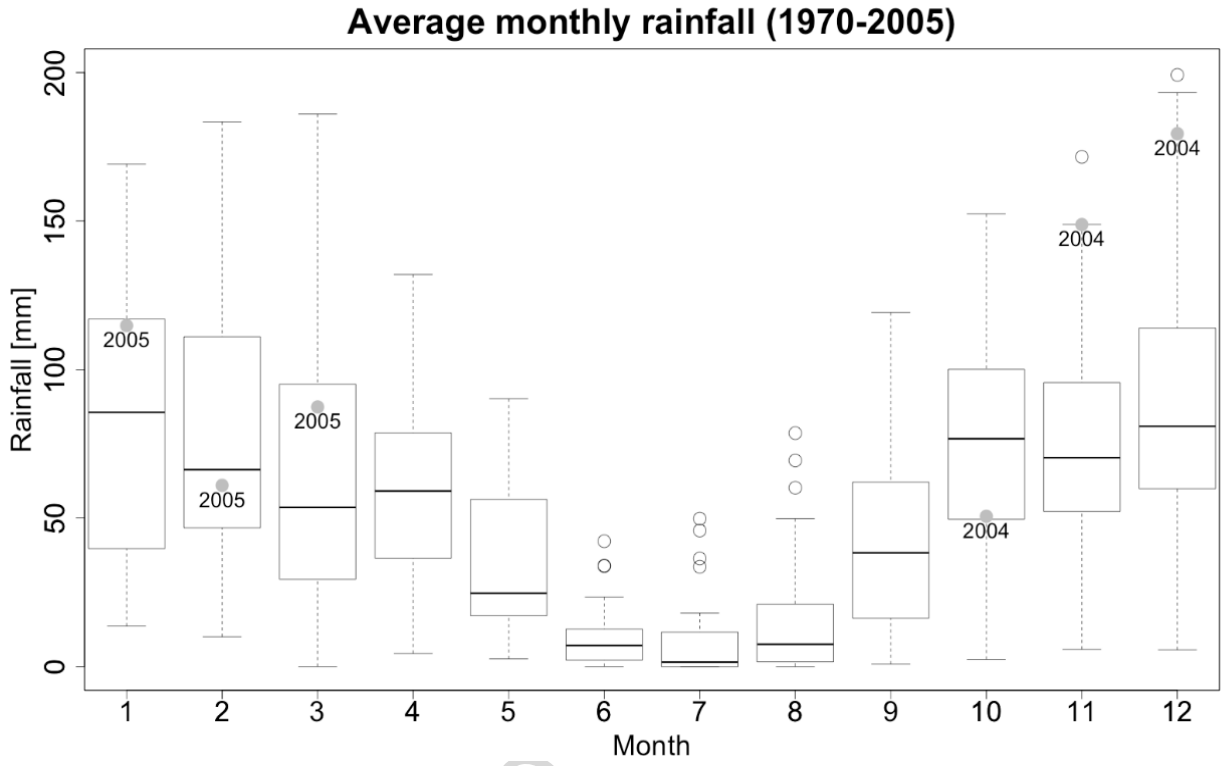


Figure 4

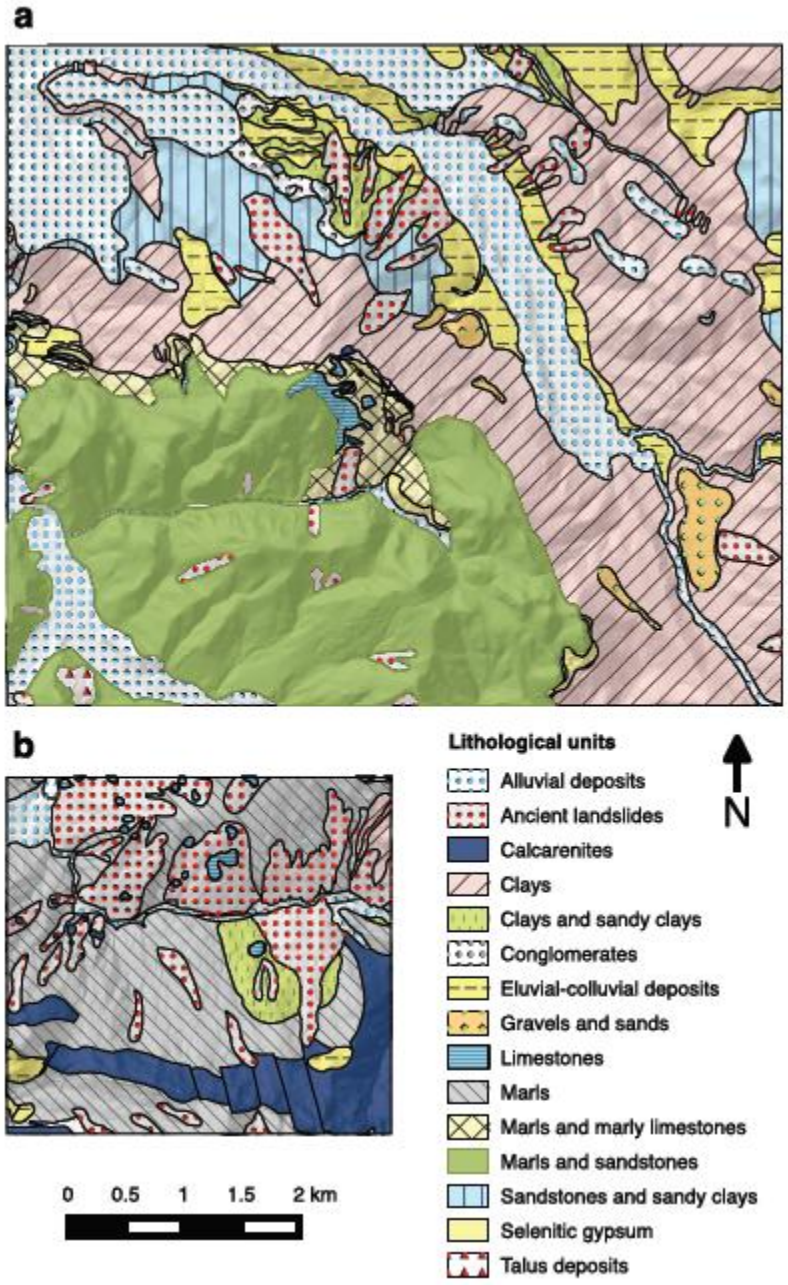


Figure 5

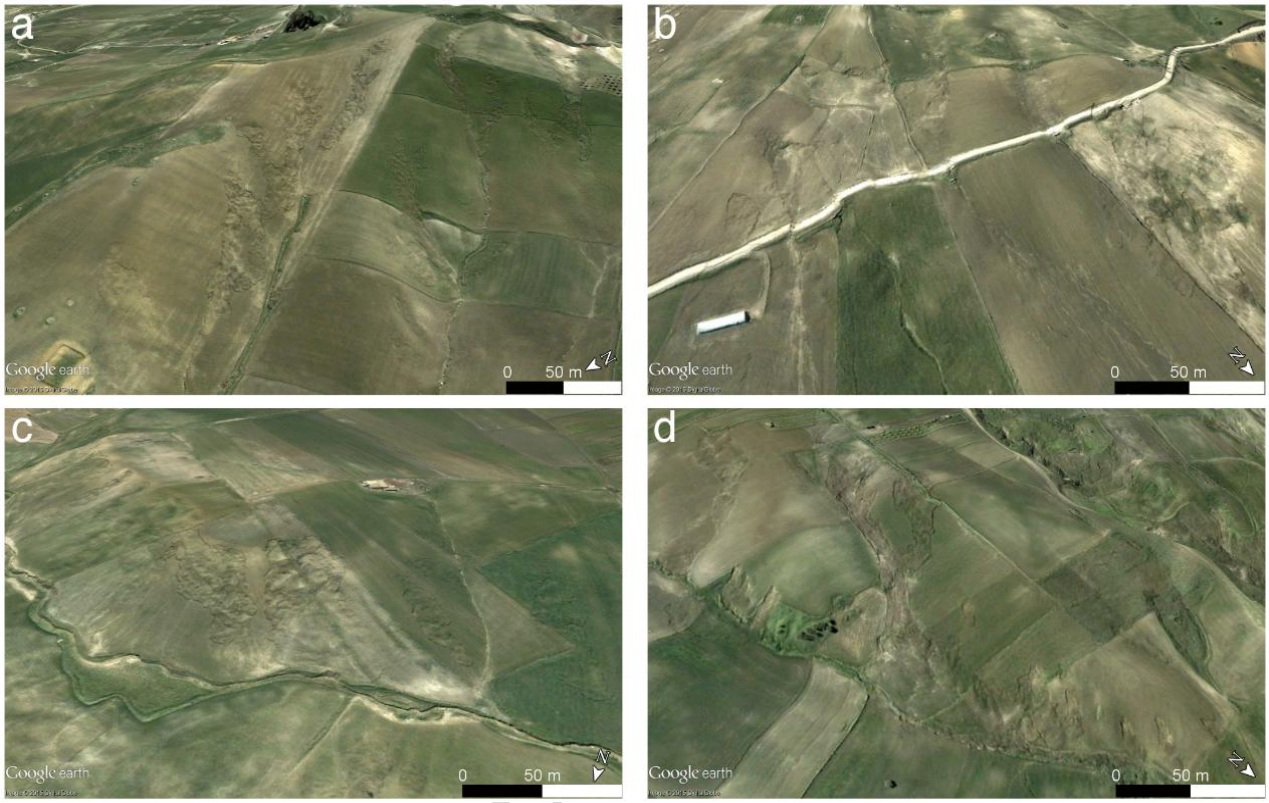


Figure 6

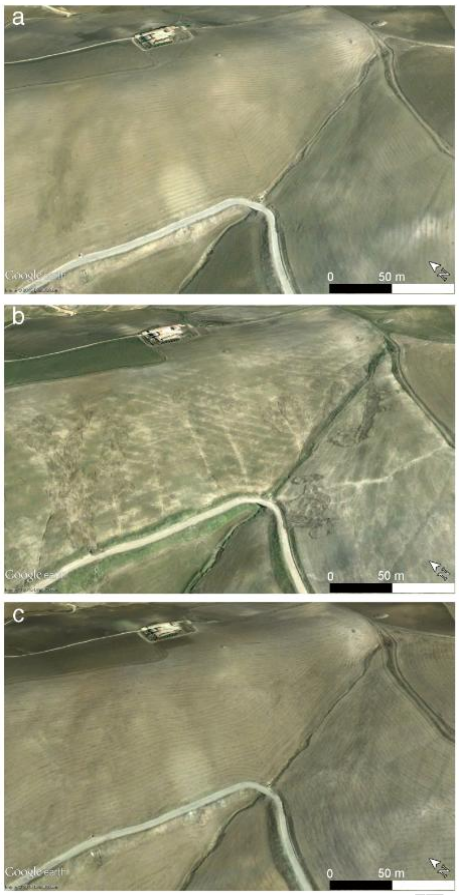


Figure 7

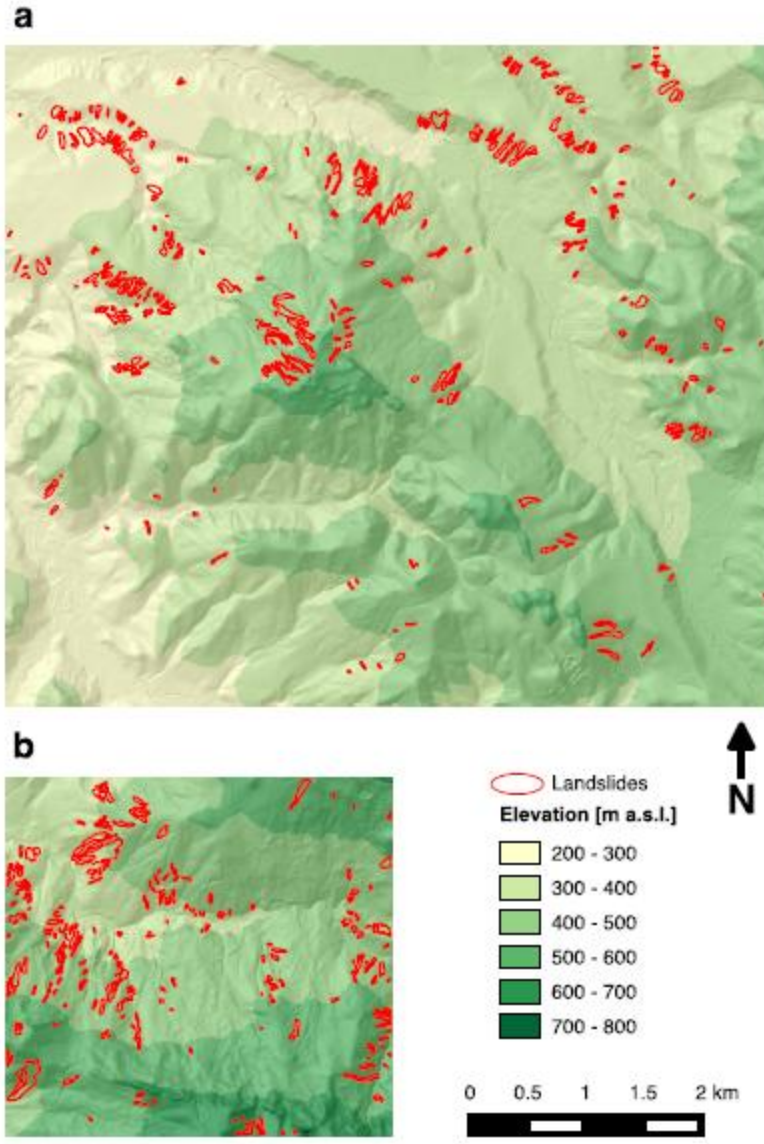


Figure 8

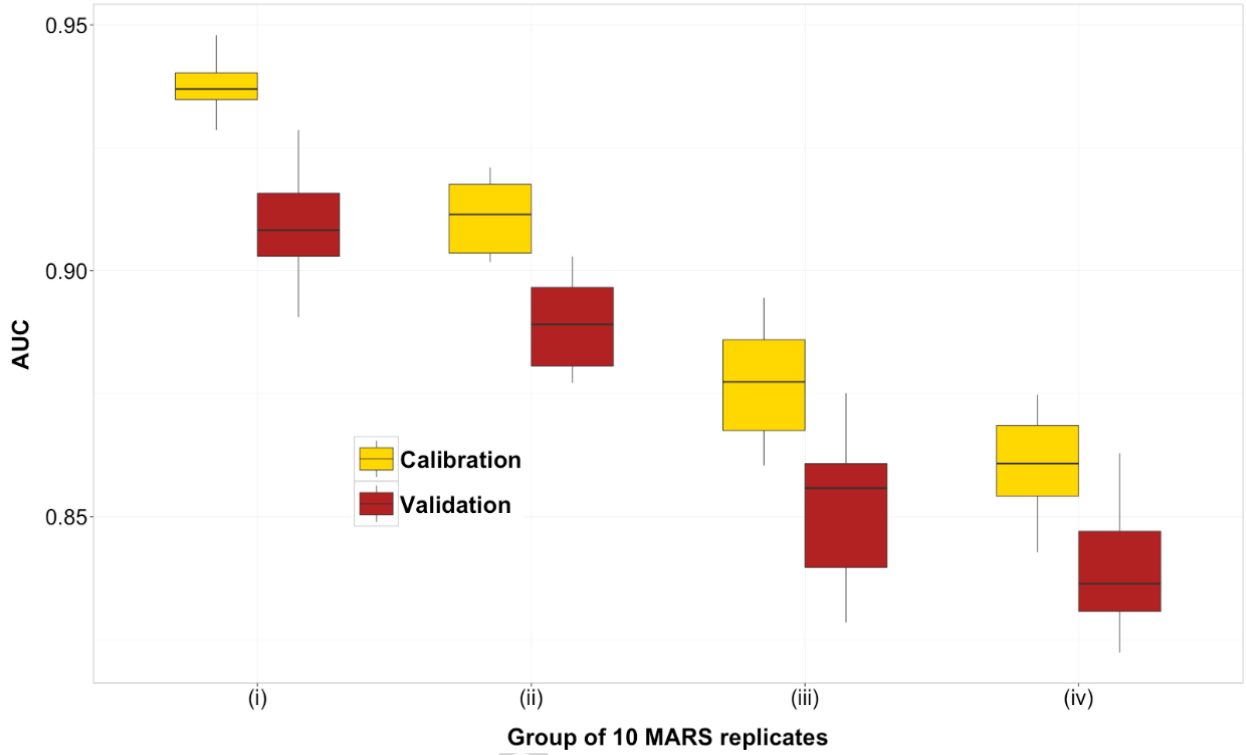


Figure 9

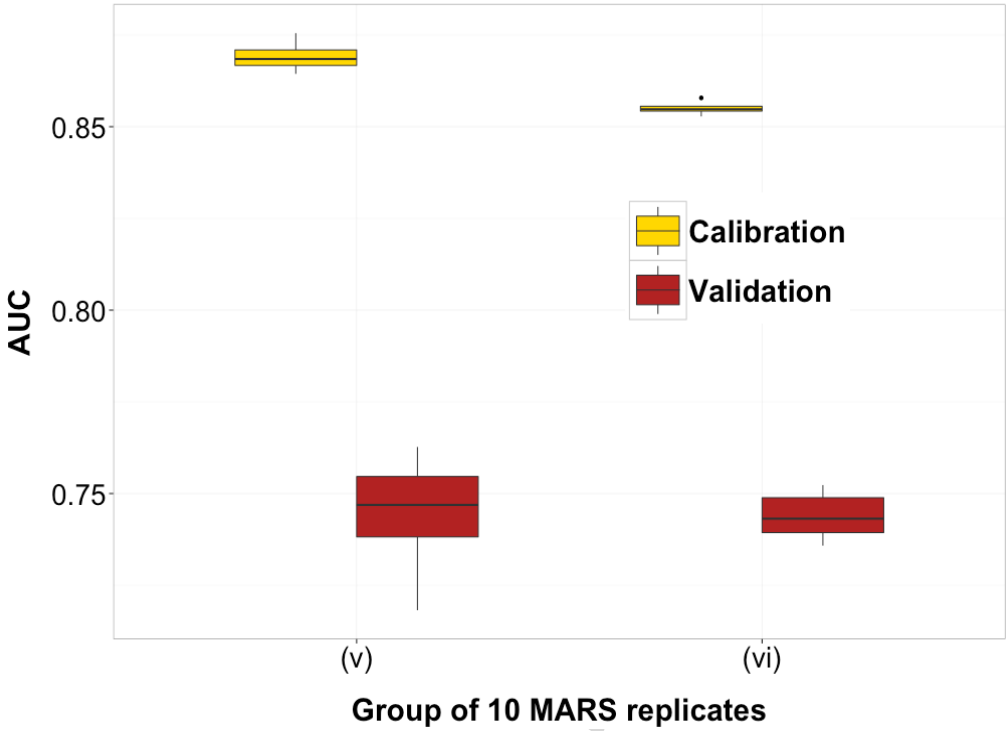
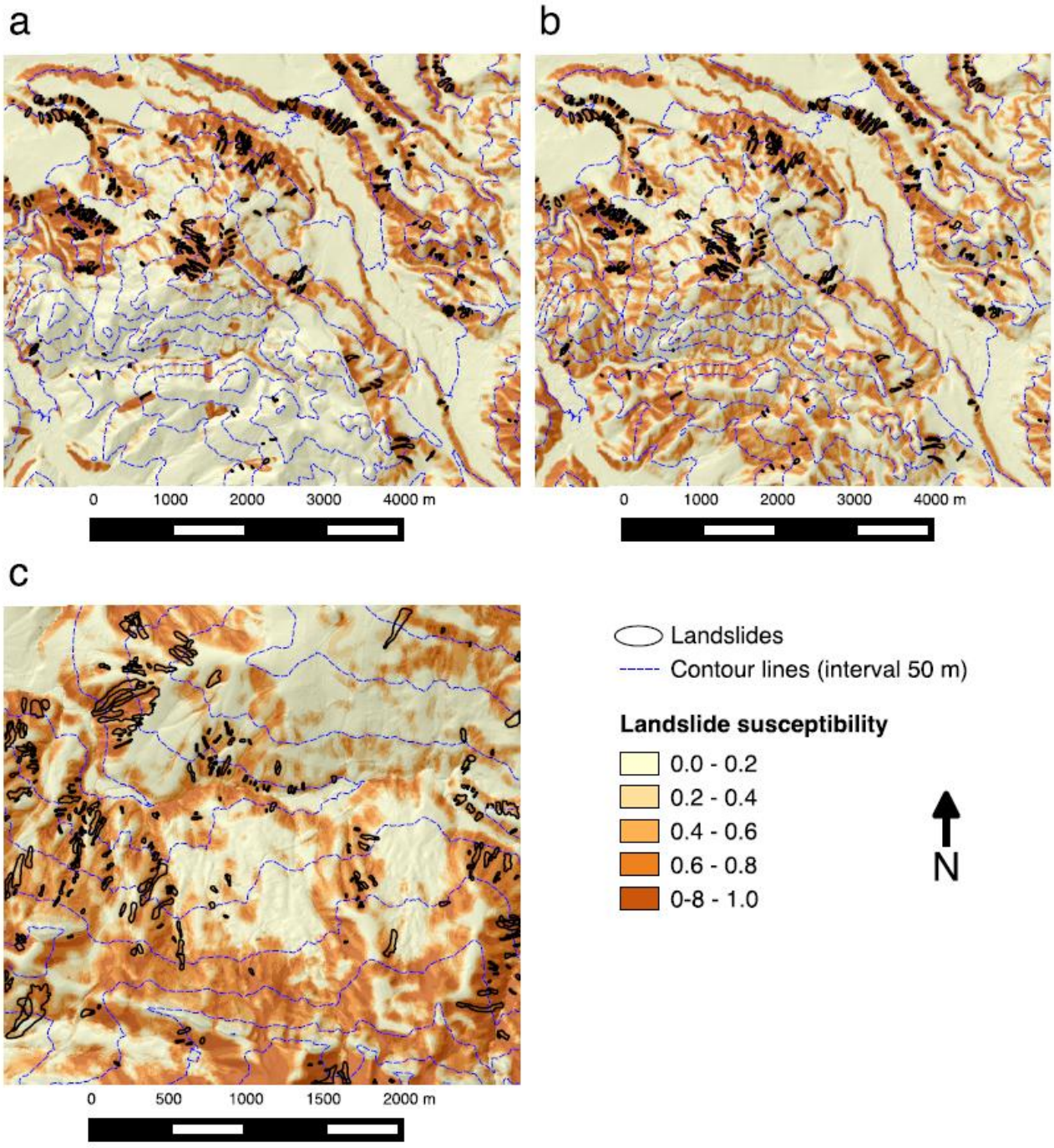


Figure 10



Highlights

- A landslide inventory was produced exclusively by visual analysis of Google Earth images.
- Landslide susceptibility models were calibrated and validated in two study areas.
- Multivariate Adaptive Regression Splines (MARS) was used as modeling technique.
- Two different methods of negative selection were tested.
- The landslide susceptibility models show acceptable to outstanding accuracy.