Nat. Hazards Earth Syst. Sci. Discuss., 3, 1731–1774, 2015
www.nat-hazards-earth-syst-sci-discuss.net/3/1731/2015/
doi:10.5194/nhessd-3-1731-2015

Natural Hazards and Earth System Sciences
Discussions

Predicting storm triggered debris flow events: application to the 2009 Ionian-Peloritan disaster (Sicily, Italy)

M. Cama, L. Lombardo, C. Conoscenti, and R. Rotigliano

Dipartimento di Scienze della Terra e del Mare, Università degli Studi di Palermo, Palermo, Italy

Received: 2 February 2015 – Accepted: 13 February 2015 – Published: 3 March 2015

Correspondence to: M. Cama (mariaelena.cama@unipa.it)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

The main assumption on which landslide susceptibility assessment by means of stochastic modelling lays is that the past is the key to the future. As a consequence, a stochastic model able to classify a past known landslide scenario should be able to predict a future unknown one as well. However, storm triggered landslide events in the Mediterranean region could pose some limits on the operative validity of such expectation, as they typically result by a randomness in time recurrence and magnitude. This is the case of the 2007/09 couple of storm events, which recently hit north-eastern Sicily resulting in largely different disaster scenarios.

The purpose of this study is to test whether a susceptibility model based on stepwise binary logistic regression is able to predict a storm triggered debris flow scenario. The study area is the small catchment of the Itala torrent (10 km^2), which drains from the southern Peloritan Mountains eastward to the Ionian sea, in the province of the Messina territory (Sicily, Italy). The shallow landslides activated in the occasion of two close intense rainfall events have been mapped by integrating remote and field surveys, producing two event inventories which include 73 landslides, activated in 2007, and 616 landslides, triggered by the 2009 storm. The set of predictors were derived from a 2 m cell digital elevation model and a 1 : 50 000 scale geologic map. The topic of the research was explored by performing two types of validation procedures: self-validation, based on the random partition of each event inventory and chrono-validation, based on the time partition of the landslide inventory. It was therefore possible to analyse and compare the performances both of the 2007-calibrated model in predicting the 2009 landslides (forward chronovalidation) and vice versa of the 2009-calibrated model in predicting the 2007 landslides (backward chronovalidation).

Both the two predictions resulted in largely acceptable performances, in terms of fitting, skill and reliability. However, a loss of performance and differences in the selected predictors between the self-validated and the chrono-validated models which are linked to the characteristics of the two triggering storms are highlighted.

1 Introduction

Landslide susceptibility is the likelihood of a landslide occurring in an area on the basis of local terrain conditions (Brabb, 1984). This definition limits the task of the prediction procedures at estimating the spatial probability for future landslides. In fact, unlike hazard assessment, susceptibility studies are only aimed to determine “where” new landslides are more likely to occur, without considering any energy (magnitude) and temporal probability (time recurrence) estimation. In spite of its more limited predictive meaning, landslide susceptibility is actually the most largely pursued task in the field of regional geo-hydrological risk assessment, both for civil protection and land use planning. In fact, knowing where a landslide is more likely to occur allows the user to define mitigation planes for non-displaceable infrastructures (e.g., roads and buildings) or to modulate the territorial vulnerability with respect to the geomorphological threats scenario in land management plan. Moreover, assessing landslide hazard with good precision and accuracy frequently requires time/money costs which are unreasonable and unbearable in regional or basin scale studies.

Landslide susceptibility assessment can be achieved by means of different methods, among which the stochastic approach has gained more and more importance in the last two decades in regional assessment applications. In fact, they produce objective, quantitative and verifiable estimates of the spatial probability for new landslides in a given study area. Moreover, stochastic approaches are very easily implementable on Geographic Informative Systems (GIS) so to exploit the more and more diffused databanks of physical-environmental attributes layers. These methods are based on some generally accepted assumptions, the basic one of which being that *the past is the key to the future* (Carrara et al., 1995). So, a susceptibility model trained in reproducing a past known landslide spatial distribution, will be able also to predict the future locations of the new failures.

For a given study area, statistical techniques allow us to derive and test for significance the multivariate relationships between the spatial distributions of an inventory of landslides (the *known target pattern*) and a set of physical-environmental variables (the predictors), which are supposed to drive the slope failures on the basis of a geomorphological model, acting as controlling factors. In the framework of the above recalled principle, the new landslides (the outcome) will occur under the same conditions which explain the present landslide distribution. Thus, a trained predictive model optimizes the functional relationships between predictors and outcome, maximizes its skill in reproducing the known target pattern (the training dataset), and it is finally tested in correctly classifying the unknown target pattern (the test dataset).

As the controlling factors are selected among the preparatory causes, which are time invariant, no matter the age of the landslide inventory exploited to train the model, as far as the basic assumption holds, any calibrated model will be able to predict any past or future unknown target pattern.

Unfortunately, very often, susceptibility assessment studies are affected by a lack of temporal information on landslide inventory which makes impossible to perform a pure chrono-validation.

Based on the scheme described above, in order to elude the lack of temporal information, strategies for the validation of the predictive models can be defined. In particular, when seasonal or event inventories are not available a validation can be performed by following a *random time partition* procedure (Chung and Fabbri, 2003). In this case, the source inventory is split into a training and a test subset to simulate the known and the unknown target patterns, respectively. In this work the above scheme is defined as a self-validation procedure to stress the circumstance that, under a morphodynamic perspective, training and test patterns actually are two partial sides of the same event. On the other hand, the term chrono-validation will be used when referring to pure temporal validation, i.e. when the training and the test target patterns are two temporally defined datasets.

It is evident how the whole scheme of the stochastic approaches is strictly dependent on the holding of the basic assumption. Any changes in the real relationships between preparatory causes and landslide activity will affect the prediction skill of the obtained susceptibility models.

Extreme events produce those morphodynamic responses that can escape from the general rule. In fact, due to intense triggering, such as a storm, the same area can result in an "out-of-range" slope response which could not be correctly predicted by a model skilled in reproducing "normal" landslide scenarios. That could resides in the non-linearity of the relationship between preparatory causes and landslides in the domain of the trigger intensity. It is therefore necessary to check for this kind of behaviour so to find a strategy which maximizes the ability of a susceptibility model to predict extreme events.

A contribution to this task is here given, exploiting a case study in north-eastern Sicily, where two recent storm events (2007 and 2009) hit the Ionian side of the Peloritani Mountains (Fig. 1). In particular, the study area is the Itala catchment (nearly 10 km^2), which is located in the southern sector of the Peloritan ridge.

In order to investigate our topic, the shallow landslides activated in the occasion of two different extreme events have been mapped by integrating remote and field surveys. A simple set of predictors was used by exploiting a 1 : 50 000 scale geological map and a 2 m cell digital elevation model (DEM). Statistical models have been obtained by applying the stepwise (forward) binary logistic regression technique (Hosmer and Lemeshow, 2000), which has been largely adopted in landslide susceptibility studies (Atkinson et al., 1998; Olhmacher and Davis, 2003; Süzen and Doyuran, 2004; Brenning et al., 2005; Carrara et al., 2008; Costanzo et al., 2014; Lombardo et al., 2014; Heckmann et al., 2014) demonstrating a great suitability to the geomorphological task and a very high performance also in comparative studies (Guzzetti et al., 2005; Rossi et al., 2010).

Exploiting a multi-temporal high resolution dataset (provided by A.R.T.A. – Assessorato Regionale Territorio e Ambiente) two landslide event inventories have been prepared allowing us to perform and validate two types of modelling procedure: self-validation, based on the random partition into a training and a test subsets of each event inventory and chrono-validation, based on the temporal partition into the 2007 and 2009 cases. The latter procedure was applied to analyse the performances both of the 2007-calibrated model in predicting the 2009 landslides (forward chrono-validation) and of the 2009-calibrated model in predicting the 2007 landslides (backward chrono-validation). By analysing and comparing the performances of the four kinds of models, the problems in predicting storm-triggered landslides are outlined and discussed.

2 General framework

2.1 Study area

The study area is located in the north-easternmost edge of Sicily (southern Italy), on the Ionian slopes of the Peloritan ridge, 20 km southward from the town of Messina (Fig. 1a). In particular, the Itala catchment is located in the Itala municipality territory, stretching for 10 km^2 and draining south-eastward for near 6 km from Mt. Scuderi (1259 m a.s.l.) to the Ionian Sea. Geologically, the area is situated between the Mandanici, Mela and Aspromonte structural units (Messina et al., 2004), which are separated by thrusts and further fractured by the neo-tectonic faults. These units are made up of high to medium grade metamorphic rocks. In particular, the Mandanici unit is mainly characterized by the outcropping of phyllites, while Mela and Aspromonte units mainly consist in paragneiss and micashists (Fig. 1b).

According to the Köppen classification, the climate in the region is classified as a Mediterranean (Csa) type, being therefore characterized by a dry season from April to September and a wet season from September to March, with an average yearly rainfall of nearly (900 mm). Besides, due to the warm water of the Mediterranean Sea and the closeness of the ridge to the seacoast, storm events are frequent in the autumn season in this sector of Sicily.

Due to their shortness and steepness, although the Ionian Peloritan torrents are usually almost dry, under raining conditions, the discharge rapidly increases some-times determining floods which affect the infrastructures (especially roads) located in the proximity of the riverbanks. Moreover, under autumn storm events, the combination of their hydrologic and geomorphologic setting occasionally determines severe morphodynamic responses, including multiple debris flow and debris flood events, such as those occurred in 2007 and 2009. The potential occurrence of this kind of events makes the whole set of Ionian Peloritan catchments, as one of the most exposed zones to hydrogeological risk in Sicily.

The inhabited areas of the Itala catchment are located in very dangerous sectors either at the base of very steep terraced slopes or near to the outlet of the streams. With respect to the land use, the area can be divided in an eastern and a western sector. The first is highly terraced and mainly cultivated with citrus groves; the second is characterised by chestnut forests and pasture characterize. The study area is strongly affected by wildfires during the summer season; this influences the density of vegetation, the soil structure and the erosional processes acting on the slopes.

2.2 The 2007 and 2009 events

The present research is based on the comparison of the landslide scenarios produced by two recent storm events which affected the Ionian Peloritan area in 2007 and 2009. The two storm events have been registered by the regional rain gauge network (Osservatorio delle Acque Sicilia) at the two stations located in Briga and Messina (Fig. 1a). The data show that on 2007 (Fig. 2) the main event (registered at Briga with 102 mm of rain in 24 h) was anticipated by longer and more extended raining periods, which lasted from the 20 to the 23 October. The storm triggered hundreds of debris flows in the whole area and 73 in the Itala catchment. Therefore, although a marked morphodynamic response was observed, limited damages and no life losses were recorded.

On the 1 October 2009 (Fig. 3) a higher magnitude event both in terms of rain quantity and number of triggered landslides hit the same area. The cumulative daily rainfall of this storm was of nearly 220 mm and, also in this case, the main rainfall event followed two previous ones (16 September: 76 mm, in six hours; 23–24 September:

190 mm, in ten hours). As a consequence of the rainfall, 616 debris flows triggered within less than five hours in Itala catchment, which produced large damages to buildings and main roads.

The typologies of the landslides which activated in the two occasions are mainly classified as channelized debris flows and debris avalanches or hillslope debris flow (Varnes, 1978; Hutchinson, 1988; Hungr et al., 2001, 2014), involving the weathered mantle of the metamorphic bedrock which outcrops over the very steep slopes of the Itala catchment. The channelized debris flows (Fig. 4a) had energy enough to reach the main river network, the discharge was extremely high and the debris was transported for long distances to the coast forming a fan. The debris avalanches (Fig. 4b) being also characterized by high energy, reached the foot of the slopes causing damages to structures and roads.

However, as the aim of the study was to study the susceptibility for new activation, the whole set of phenomena was processed as a single type, using in the following the general sense of the term *debris flow*. The very few cases of bedrock-landslides, such as falls and rotational slides, were deliberately excluded from the analysis, as they would have required a different approach both in terms of controlling factors and statistical methods.

In both the cases, the main events were anticipated by fore-storm rainfall events, which had been responsible for the saturation of the weathered mantle of the metamorphic bedrock (Aronica et al., 2012). Therefore, when the main storms hit the area, the high water content very rapidly reduced the shear strength of the regolithic layer determining the contemporaneous activation of debris flow.

The difference in magnitude between the two phenomena, was not only related to the number of activated phenomena but also in the kinematic behaviour of the landslides. In fact, on the 2007, the percentage of phenomena which reached the foot of the slope, or the main channel, was lower. This suggests that the quantity of water was enough to saturate the soil and trigger the shallow landslides but not enough to determine a long

NHESSD
3, 1731–1774, 2015

distance transport. On the contrary, on the 1 October 2009 the slope conditions and the consequences were quite different, especially in terms of damages.

A number of studies have been recently published regarding the 2009 event, but mainly focused on the northern catchments of Giampilieri and Briga torrents (Aronica et al., 2012; Del Ventisette et al., 2012; Lombardo et al., 2014), on the techniques for landslide recognition (Ardizzone et al., 2012) and on the influence of land use in landslide susceptibility (Reichenbach et al., 2014). No focus was made on problems in chrono-validation, which could play in the opinion of the Authors a key point in estimating the reliability of the susceptibility models in terms of future predictability of new events.

3 Materials and methods

Among the large set of statistical methods which are capable to fit those functional relationships which link the probability of an outcome to a set of predictors, in the last two decades binary logistic regression (BLR) has became one of the most applied methods in landslide susceptibility studies. In fact, it does not requires heavy constraints on the statistical distributions of the predictors and allows the user to include in the model both nominal and continuous variables. The structure of a BLR model is very simple and geomorphologically interpretable, being composed of single coefficients which describe the linear correlations between each predictors and the log-odd of the binary outcome (stable/unstable).

The application of BLR for landslide susceptibility assessment typically requires the following steps: the partition of the study area into mapping units, which are then characterized with respect to a set of potential predictors; the assignment to each mapping unit of its stability conditions, based on its spatial relationships with past landslides (e.g., inclusion or intersection); the extraction of a balanced (stable/unstable) dataset from the whole set of mapping units; regression of the modelling function; validation of the model.

NHESSD
3, 1731–1774, 2015

This chapter describes the methods and the model building strategy which have been adopted to investigate the main topic of the research: exploring skills and limits in predicting a storm triggered multiple debris flow event.

3.1 Landslide inventory

Landslide recognition was performed by integrating a field survey, which was carried out soon after the 2009 disaster, and orthophoto analysis which allowed to visualize the slopes at different dates. In particular, the high resolution LiDAR (Light Detention And Ranging) data from two different acquisitions, respectively 2008 and 2009, were used. These data were provided by the Territory and Environment Department of the Sicilian Government (ATA2008 – Assessorato Regionale Territorio e Ambiente) and the National Civil Protection (PCN2009, Protezione Civile Nazionale). The ATA2008 data (taken in August) includes 0.25 m pixel orthophotos and a DEM having 2 m and 0.22 for horizontal and vertical resolution, respectively. The PCN2009 data were acquired six days after the 2009 event and includes 15 cm pixel orthophotos, and a 1.1 m cell DEM. Besides, multi-temporal (2005, 2006, 2010 and 2012) Google Earth™ (GE) images were analysed so to compare the 2007 and 2009 mapped phenomena with the previous and successive slope conditions.

An event inventory (Guzzetti et al., 2012) has to report only those landslides which have been triggered by a single specific trigger occurrence, such as an earthquake, rainfall or snowmelt. To fit this constraint, a first landslide mapping was carried out on the 2008 and the 2009 images, comparing also their morphological signature on the correspondent hillshaded DEM and obtaining a first version of the 2007 and 2009 inventories. The mapped phenomena were supposed to be activated during the 26 October 2007 for the first inventory and on the 1 October 2009 on the second. While for the 2009 this assumption is quite verified because the LIDAR data were acquired only 5 days after the event, for the 2007 an historical check of heavy raining events and associate landslide activity was necessary. Then, the morphologies mapped on 2007 were compared with the 2006 GE-images. By combining the data obtained from the

NHESSD
3, 1731–1774, 2015

three time frames, five different cases were obtained (Fig. 5): (a) shallow landslides mapped on the 2007 orthophoto but activated before the 2007 event; (b) shallow activated during the 2007 event which did not reactivated or retreated during the 2009; (c) shallow landslides activated during the 2007 that retreated or reactivated during the 2009; (d) shallow landslides activated during the 2007 which have been completely eroded during the propagation phase of the 2009; (e) shallow landslides which activated during the 2009 event in precedent stable areas.

The final event inventories (Fig. 6) contain 73 events for the 2007, corresponding to cases b, c and d, and 616 for the 2009, corresponding to the case e. Each landslide inventory was stored into two separated vector layers: the first containing a polygon representing the source areas, the second containing the Landslide Identification Points, corresponding to the highest point along the crown of each mapped phenomenon (LIP, Costanzo et al., 2012a, 2014; Lombardo et al., 2014).

3.2 Binary logistic regression

Binary logistic regression (BLR) is a multivariate statistical technique, based on a frequentist approach, which is used to model the expected value of a response variable (the outcome) by a linear combination of either continuous and discrete predictor variables (Hosmer and Lemeshow, 2000). With respect to other frequentist methods (e.g., the discriminant analysis), it does not require any linearization or transformations to obtain normal distributed covariates. Moreover, the outcome of BLR is easily interpretable for applied scientists.

In binary logistic regression the response variable Y assumes one of the two mutually exclusive values of 0 (no landslide) or 1 (landslide) for stable mapping units or unstable mapping units, respectively.

The relationship between the predictors and the probability for the response variable to assume the value 1 is linearized by the logit function logit(Y), which corresponds to

NHESSD
3, 1731–1774, 2015

the following transformation:

$$\text{logit}(Y) = \ln[P(Y = 1)/(1 - P(Y = 1))] = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n; \tag{1}$$

where $P(Y = 1)$ is the probability that the response variables assumes the value 1, α a constant term or intercept, the $x_1, x_2, \ldots x_n$ are the input predictor variables and the β_n their coefficients. Therefore, once the logit function is calculated, and the β_n values are known, the probability can be back-calculated using the following formula:

$$P(Y = 1) = e^{\text{logit}(Y)}/[1 + e^{\text{logit}(Y)}]; \tag{2}$$

This equation ensures that, for any given case, the probability $P(Y = 1)$ will not be less than 0 or greater than 1 with $\text{logit}(Y) = \pm\infty$.

The sign of the β coefficients joined to the odd ratios (OR), which is calculated by simply exponentiating the β, indicates how likely (or unlikely) it is for the outcome to be positive (unstable cell) when a unit change of an independent variable occurs. Negatively correlated variables will produce negative β s and OR limited between 0 and 1; positively correlated variables will result in positive β and OR greater than 1.

In order to estimate the best intercept and βn coefficients, the logistic regression uses the maximum likelihood technique. This approach maximizes the value of the log-likelihood function (LL), which indicates how likely is to obtain the observed value of Y, given the values of independent variables and coefficients (Menard, 2002) In particular, the global fitting of the regressed model on the data domain is usually expressed by the −2LL (negative log-likelihood) which is an estimator based on the maximum likelihood criterion. The differences in −2LL value between the model with only the intercept ($L_{\text{INTERCEPT}}$) and the full model (L_{MODEL}) have a χ^2 distribution, so that the significance of the regressed coefficients can be easily tested (Olmacher and Davis, 2003; Akgun and Turk, 2011). In other words, the −2LL test estimates the significance of the increase in model fitting produced by the introduction of the predictors.

In the present research, we applied BLR under a stepwise selection routine, which was already successfully adopted in landslides and debris flows susceptibility studies (Begueria, 2006; Atkinson and Massari, 2011; Meusburger and Alewell, 2009;

Costanzo et al., 2014; Heckmann et al., 2014; Lombardo et al., 2014). The stepwise selection is an iterative procedure, which selects the best performing and most parsimonious set of predicting variables. It can be performed either in forward or in backward mode. In the first case the procedure starts from an "intercept only" model and consists in selecting and adding at each step, from the group of available variables, the one which results in the larger likelihood increase. On the contrary, the backward stepwise selection starts from a full model including all the variables and removes iteratively the variables until the model reaches the best fitting.

In the forward stepwise selection the model global fitting at each stage of the iterative process is determined by the Akaike information criterion (AIC; Akaike, 1973). More in detail, the AIC value increases for the number of variables and decreases with a larger likelihood function.

At every step the procedure introduces iteratively all the variables and selects the one that minimized the AIC and maximized the -2LL values. The first factor is the one which produces the greatest change in the log-likelihood, with respect to the intercept. The iterative calculation stops when the addition of any of the left variables does not meaningful increase the performance of the model. The result is the restricted list of variables, each having its order of importance (i.e. the iteration in which it was picked up) that can be submitted to the final BLR.

All the statistical analyses which are hereafter discussed were performed by using an open source software (TANAGRA: Rakotomalala, 2005).

3.3 Covariates and outcome status assignment

The first step in modelling the landslide susceptibility using a stochastic approach is to select those mapping units in which the study area has to be partitioned. Mapping units are the basic spatial elements in which the model will be able to produce a prediction. Two main types of mapping units are adopted in literature: hydro-geomorphological units and regular grids. The former allows the model to exploit the morphodynamic homogeneity of the area which is included into each single unit, corresponding to hydrological or slope units; the latter optimize the matching between the spatial resolution of the source layers of some important predictors, typically having the same grid structure of the DEM. Moreover, in debris flow susceptibility studies the use a regular grid simplifies the selection of potential source areas for propagation modelling.

Therefore, selecting a mapping unit means defining the topological structure of the model, which can be either randomly vector defined or regularly raster based.

In the present research a raster-based structure was adopted by partitioning the study area into a 8 m square cells grid, which required the rasterisation of the spatial distribution of all the covariates no matter their source structure.

Starting from a DEM and a geological map, the following eight potential predictors have been selected so to assign their value to each cell in which the study area has been partitioned (Figs. 8 and 9): Outcropping lithology (GEO), Land use (USE), Aspect (ASP), Steepness (SLO), Topographic Wetness Index (TWI), Plan (PLAN) and Profile (PROF) curvatures and Distance from tectonic features (DFAULT).

Outcropping lithology and tectonic features are proxy variables expressing the mechanical properties of the bedrock and the weathered mantle. These variables were obtained from a 1 : 50 000 available geological map (Lentini et al., 2007), which was derived from 1 : 10 000 field surveys.

Land use allows the model to summarize those potential modifications of the natural structure of the regolithic mantle and the bedrock, which are related to anthropogenic activities. In order to adapt the resolution of this a land use map, based on the analysis of the orthophotos ARTA2007/2008 and PCN2009 and field recognition was prepared. The final land use map contains 6 classes: (i) medium-high vegetated terraces (MHVT); (ii) low vegetated terraces (LVT); (iii) chestnut forests (CF); (iv) pastures (P); (v) urbanized areas (UA); (vi) river beds and beaches (RB).

Slope steepness, Plan and Profile curvatures are related with the energy of the relief. Steepness is commonly used as predictor in landslide susceptibility and very often it presents a very high importance. In fact, especially for debris flow analysis it is expected to be one of the most significant variables because it is directly linked to the

NHESSD
3, 1731–1774, 2015

shear strength acting onto the potential shallow failure surface. Moreover, for shallow failures presenting slide or flow mechanisms, the topographic surface and the rupture plane or zone can be considered as almost parallel. In this case, the slope steepness is a proxy for the real inclination of the potential failure surface. Steepness also controls the overland and subsurface flow velocity and runoff rate. At the same time, the topographic curvatures control the divergence and convergence, both of surface runoff and shallow gravitational stresses (Ohlmacher, 2007). In this study the *profile curvature* and the *plan curvature* were used, which correspond to the second derivatives of the slope steepness and the aspect, respectively. Curvatures are expected to be the best proxy variables for convergent flow of water (plan curvature) and changes in flow velocity (profile curvature).

Topographic Wetness Index is defined as $\ln(\mathrm{As}/\tan\beta)$ where As is the local upslope area draining per contour unit length and β is the local slope angle. It describes the extension and distribution of the saturation zones assuming steady-state conditions and uniform soil properties (Moore et al., 1993), demonstrated through the comparison of field data, that TWI can be considered a proxy variable directly related with the properties of soil, in particular with the soil moisture, A horizon depth, Phosphorus content and organic matter.

Aspect controls the intensity at the earth surface of the solar insolation, and as a consequence, the evapotranspiration and flora and fauna distribution and abundance. Being the erosional processes related with the chemical physical weathering operated by water, temperature and vegetation, it is very important to consider this factor for the determination of landslide susceptibility. Besides, ASP frequently assumes a role of proxy variable for the attitude of the rock layers.

The input for the calculation of the topographic attributes was the DEM ARTA 2007/2008 resampled at 8 m pixel size with the nearest neighbour approach. The resampling operation on the original DEM (2 m pixel size) smoothed the effects of microtopography and possible noise existing on the original data.

All the factors have been calculated using SAGA GIS (System for Automated Geoscientific Analysis, Conrad 2007).

Once the 8 m grid layers of the predictors were obtained, they were combined in a single multivariate one, which was crossed to the LIP vector layers, to set the stable/unstable status. Each cell hosting at least one LIP was set as unstable.

3.4 Validation procedures and model building strategy

Model validation is a mandatory component of a susceptibility assessment studies (Carrara et al., 2003; Guzzetti et al., 2006; Frattini et al., 2010; Rossi et al., 2010). No matter the method adopted in modelling the susceptibility, rigorous and quantitative validation procedures furnish the only criterion for accepting or rejecting a predictive model.

The validation of a model requires the availability of a training and a test set of landslides or outcomes. The training landslides are exploited to calibrate the maximum-likelihood fitting, so to optimize the regression coefficients; the predicted probability which is generated by the model is then compared to the actual target pattern which is defined by the test landslides set. The accuracy of a model is then evaluated by comparing the produced prediction image to the known (training) and unknown (test) target patterns. In particular, the degree of fit expresses the ability of the model to classify the known cases, while the prediction skill is the ability to predict the unknown cases.

Training and test landslides can be obtained by Chung and Fabbri (2003): time partition, random time partition or spatial partition. The first is possible when multi-temporal landslides inventories are available, the second is based on randomly partitioning single-epoch datasets and the third on sub-dividing the study area in two similar sub-sectors. Random time partition procedures can be applied either on the landslide inventory (Conoscenti et al., 2008a) or on the mapping units database (Conoscenti et al., 2008b), whilst spatial partition can also be performed also on not nested or adjacent areas such as in the study aimed at susceptibility model exportations (von Ruette et al., 2011; Costanzo et al., 2012b; Lombardo et al., 2014).

However, validating a model requires for testing its accuracy, precision, robustness and geomorphological adequacy or coherence, both in terms of predictive performance and inner structure of the model, the latter corresponding in a stepwise BLR procedure, to the rank and the coefficients of the selected predictors (Frattini et al., 2010; Costanzo et al., 2014; Lombardo et al., 2014). Besides, as BLR requires for balanced (positive/negative cases) datasets, a single regressed dataset has to contain the positive cases (unstable cells) and an equal number of randomly selected negatives (Atkinson et al., 1998; Süzen and Doyuran, 2004; Nefeslioglu et al., 2008; Bai et al., 2010; Costanzo et al., 2014; Frattini et al., 2010; Van Den Eeckhaut et al., 2009), which could determine a low representativeness of the analysed cases. In particular, in this study, each pixel containing a LIP has been considered as diagnostic area (Rotigliano et al., 2011), while the negative cases have been randomly selected in the catchment, outside the landslide polygons. In order to obtain a better dispersion of points and to avoid autocorrelation of the spatial variables, the distance in the random selection was maximized. Therefore, every model was optimized by regressing 10 datasets, each containing 146 balanced cases (positive/negative), for 2007, and 1232 balanced cases, for 2009. This heavily reduces the number of really analysed cases to a very small percentage of the cells in which the study area is partitioned, so that a need of testing the representativeness of the worked subset also arises. To control the possible effects introduced by this procedure, multi-extraction of negatives are to be performed and more than one dataset regressed. In particular, a multiple extraction produces m different balanced datasets, each composed by the union of the same positives and a different set of randomly extracted negatives; multi-fold cross validation procedures are then applied, by resampling n times the same dataset to perform n replicates of the regression procedure, finally obtaining $n \times m$ outcomes of the same performance indexes or model parameters.

In this research, two suites of ten dataset were extracted both for 2007- and 2009-models; to each dataset a 10-fold cross validation procedure was then applied, which gave for each mapping unit a total of one hundred probability estimates (10 replicates × 10 subsets), based on which accuracy and precision of the predictive performance were tested. Moreover, each of the one hundred replicates resulted in a set of ranked predictors and regression coefficients the comparison of which allowed us to test the precision and the robustness of the model.

Once a cut-off for the estimated probability is fixed to split positive and negative predictions, the crossing with a target pattern results in the production of true positives (TP), true negatives (TN), false positives (FP: Type I errors) and false negatives (FN: Type II errors) cases. Contingency tables are used to summarize these data and to compute the model error rate, $(TP+TN)/(FP+FN)$, sensitivity or hit rate, $(TP/(TP+FN))$, and $1-\text{specificity}$, $(FP/(TP+FN))$.

A cut-off independent technique for estimating the accuracy of a predictive model is represented by the Receiver Operating Characteristic (ROC) curves, which draw the trade-off between success and failures for decreasing probability threshold, in sensitivity $(TP/(TP+FN))$ vs. $1-\text{specificity}$ $(FP/(TP+FN))$ plots. The Area Under the Curve (AUC) in the ROC plots is the most adopted metrics for the accuracy of the predictive models.

The precision and accuracy of the model can be also represented in spatial terms, by preparing prediction and error maps, in which for each mapping unit the mean susceptibility and the dispersion of its estimates are plotted and compared to the actual distribution of the unknown positives.

In order to investigate the main topic, two kind of modelling procedures have been followed in the present research.

A self-validating scheme was applied for each of the two event-inventories (2007 and 2009), by randomly splitting (90/10 %) the 10 extracted balanced datasets of the two temporal suites in a training and a test subset. For each dataset, the random splitting procedure was applied 10 times, resulting in one hundred self-validated replicates.

A chrono-validating scheme was then applied, by training the model with the whole event-inventory of one epoch and testing the performance in matching the event-inventories of the other. We hereafter refer to forward chrono-validation, if training with 2007 and testing with 2009, and vice versa to backward chrono-validation, if training with 2009 and testing with 2007. For each temporal model suite, we produced ten prediction images based on the ten datasets of the other suite, again having one hundred backward and one hundred forward chrono-validated replicates.

4 Results

The results of the cross-validation procedures for the 2009 and 2007 self-validating one-hundred models are presented in Fig. 9. Generally, the 2009 models resulted in a more performing prediction with a lower error rate (0.33, for 2007; 0.22, for 2009) and variability. Similarly, the ROC-AUCs attest for the good quality of the models, with a better performance for the 2009 model (2009-AUC = 0.86, 2007-AUC = 0.79) and no clue of overfitting.

Once the overall quality of the predictive performance was assessed, full (without splitting in training and test subsets) regressions were run for the ten dataset of each event-inventory, so to optimize the fitting of the model and explore their inner structure.

For both the full models, the obtained ROC-AUCs are above the good performing threshold (> 0.81, 2007, > 0.87, for 2009), with error rates of 0.26, for 2007, and 0.22, for 2009.

As regards the predictors, the 2007 model suite selected 5 variables (Fig. 10), four of which with a frequency of more than 5/10: West and South-West slope aspect, steepness and FDNb outcropping lithology resulted as the main causative factors for the 2007 landslides.

A larger set of variables (17) was included by BLR in the 2009 model suite (Fig. 11), 15 of which were selected more than 5 times. Among the topographic variables the most important resulted: steepness, all the pixels without any northward aspect com-

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

NHESSD
3, 1731–1774, 2015

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◂◂ | ▸▸
◂ | ▸
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

ponent, profile curvatures (both concave and convex) and plan convex curvature of slopes. Together with topographic variables, FDNb and MLEa lithologies, distance from tectonic elements (DFAULTS) and CF (Chestnut forests) and P (Pastures) land use classes were always selected with high and stable ranking. For what concerns the β-coefficients only the variables DFAULT, concave profile curvature and the land use classes (CF and P) showed negatives value corresponding to inverse correlations with landslides.

The chrono-validated models (Fig. 12) resulted in largely acceptable ROC-AUCs (> 0.75) and error rates (< 0.3), both for forward and backward predictions. Thus, a loss in the predictive performance of both the forward and backward temporal prediction is observed. In particular, by comparing the self- and the chrono-validation performances, an AUC-decrease from 0.81 to 0.77, for 2007, and from 0.87 to 0.78, for 2009, were observed. Besides, the mean error rates increase from 0.26 to 0.30, for 2007, and from 0.20 to 0.28, for 2009. It is worth to note the more relevant performance decrease which affected the 2009 model, so that the two chrono-validations are almost equivalent. In Fig. 13, the obtained mean (over 100 replicates) ROC curves are shown.

On the basis of the obtained models, susceptibility and error maps for the 2007 and 2009 chrono-validating models were prepared (Fig. 14). In particular, the susceptibility maps show the spatial distribution of the mean (over 100 replicates) probabilities, while the error maps describe the dispersion of the estimates, represented by a 2σ interval.

At a first glance, the two susceptibility maps appear quite different: the 2007 map shows a more diffused and graduated susceptibility, with the north-western and south-eastern sectors of the catchment hosting high susceptible areas. On the contrary, the 2009 map is affected by a marked spatial separation between the north-eastern homogeneous very highly susceptible sector and the remaining larger part of the catchment, which has a very low susceptibility. In terms of error maps, the forward model is affected by a general higher level of error, with the maximum values in the central sector and very low values along the stream network. The backward model, on the contrary, produced very low errors, with the exception of the stream network, which is characterized by quite higher values, and two single small areas, corresponding to the outcropping of very poorly diffused lithologies (see Fig. 1).

To compare the two landslide susceptibility maps, also taking into consideration the distribution of the occurred landslides, 2007 and 2009 LIPs were located onto a map of the residuals, which shows the differences between the two (forward and backward) chrono-validated mean susceptibilities (Fig. 15). The map confirms the difference between the two models in estimating the susceptibility of the catchment, with higher probabilities in the southern and north-western sectors, for the forward, and in the north-eastern sector, for the backward validated models, respectively.

By comparing the two susceptibility estimates in a dispersion density plot (Fig. 16), the above-described trend is confirmed. The two models linearly agree in the higher range of susceptibility, whilst a larger dispersion exists in the lower and intermediate susceptibility range. In particular, for the stable areas (near the origin of the plot) the higher densities pixels shift toward a more than 45° steep linear trend, marking an overestimation for 2007-calibrated model. On a binarised perspective, by setting at 0.5 the cut-off value for stable/unstable discrimination, the final number of joint true predictions (II and IV sectors) is 0.77 %, no matter the calibration we adopt. False predictions (I and III sectors of the plot) resulted to reach the 0.23 %, with the two chrono-validated models performing with similar results: the backward-calibrated model fails for eleven 2007-positives (0.155 %), while the forward-calibrated misses ninety 2009-positives (0.146 %).

5 Discussion and conclusions

The main task of this research was to investigate those potential limits which could arise when analysing the predictive performances of stochastically based landslide susceptibility models. To investigate this topic, we prepared models based on two different set of debris flow events, so to train two different chrono-validated models: a forward model, trained on 2007 and tested on 2009, and a backward model, trained on

2009 and tested on 2007. Under the assumption that the past is the key to the future, the performances of the two types of chronological modelling should have been the same. However, the two landslide inventories were different in number and location in the catchment of the triggered debris flows. Therefore, we expected this could have resulted in an asymmetry in the predictive performance of the two chrono-validating models.

The analysis of the self-validated models pointed out that the 2009 model was more performing and richer in terms of selected variables. This could be interpreted as a direct consequence of the greater number of landslides which compose the 2009 inventory (one order of magnitude more), so that the fitting of the model has reached a more accurate determination of the regression coefficients. As regards the variables which were included into the two models, a different but overlapping set of factors emerged, with the 2009 model structure being richer and more articulated. In particular, slope morphology (steepness, curvature and aspect), soil use and outcropping lithology control the debris flow susceptibility in the catchment.

The comparison between the performances of self- and chrono-validating models has highlighted a loss in accuracy which was a bit more marked for the higher performing self-validated 2009 model. Therefore, if a large difference between the accuracy of the two self-validated models was observed, a very smoothed residual difference pointed out from the comparison between the forward and backward chrono-validated models. This suggests that, in spite of the higher performance which the 2009 model obtained in classifying the same 2009 event, its skill in back-predicting the 2007 landslides, was the same showed in the 2007 in forth-predicting the 2009.

On the one hand, the above-described results confirm the symmetry between forward and backward chrono-validations and the main assumption on which modelling is based. On the other hand, the loss in performance which the 2009 suffered suggests that using self-validated models for temporal prediction can mislead the user in estimating the performance of the model. In fact, one would expect that the model trained with the largest landslide inventory would be the best performing in chrono-validation as it

NHESSD
3, 1731–1774, 2015

“includes” also the less extreme responses. However, in spite of the similar overlapped inner structure of the 2007 and 2009 models, the predictive performance of the 2009-backward model lowered to the same AUC and error rates results of the 2007-backward model. The reason for this behaviour could be connected to the local characteristics of the storm events, which differently hit even a so small catchment. This would mean that inside a 10 km^2 are there are two different past and two different futures, depending on which storm event we refer to. At the same time, a non-linearity of the morphodynamic response could affect the performance in chrono-validation: a larger event does not produce a larger but a different response. The larger the difference between the triggering events, the larger the difference in the response of the preparatory conditions. This was confirmed by the different inner structures of the models.

However, from a risk prospective, the difference between the two models did not produced a relevant loss in prediction on a binary basis. In fact, only a limited number of cases resulted in a false prediction. That’s why the mapped landslides are largely located in the very susceptible pixels. However, the results of the present research confirmed that the larger difference between the two models was obtained in the intermediate susceptibilities interval, which is the same region of the error plots where the self-validated models suffer for poor precision.

In conclusion, in light of the results of this research, the direct use of “the past is the key to the future” basic assumption must be critically accepted. The holding of such a premise depends on the homogeneity of the raining events which have triggered the landslides used to calibrate the models. This homogeneity should be verified in terms of spatial distribution and intensity of the trigger. In fact, in case of differences, non-linear model effects could modify the accuracy of the models. Moreover, non-linearity effects limit the prospective of using the most severe available inventory for training the best performing model. In this research it was verified that the best performing self-validated model did not resulted as more accurate one in chrono-validation.

Storm-triggered landslide scenario are the events which typically stress the basic assumptions for stochastic modelling. At the same time, debris flow multiple scenarios are the one which more severely produce damages and life losses in several regions. Authors consider this basic topic as one of the most important, but unfortunately not considered one. Further applications to other study cases could allow the scientific community to effectively weight the accuracy of the very sophisticated statistical models which are nowadays largely adopted in landslide susceptibility assessment.

Author contributions. Authors have commonly shared all the part of the research as well as of the manuscript preparation.

Acknowledgements. The findings and discussion of this research are the results of the research activity which was carried out in the framework of the PhD research projects of Mariaelena Cama and Luigi Lombardo at the "Dipartimento di Scienze della Terra e del Mare" of the University of Palermo (XXV cycle). Luigi Lombardo PhD thesis is internationally co-tutored with the Department of Geography of the University of Tübingen (Deutschland).

This research was supported by the project SUFRA_SICILIA, funded by the ARTA-Regione Sicilia, and the FFR 2012/2013 project, funded by the University of Palermo.

References

Akaike, H.: Information Theory as an Extension of the Maximum Likelihood Principle, in: Second International Symposium on Information Theory, edited by: Petrov, B. N. and Csaki, F., Budapest, Akademiai Kiado, 267–281, 1973.

Aronica, G. T., Brigandí, G., and Morey, N.: Flash floods and debris flow in the city area of Messina, north-east part of Sicily, Italy in October 2009: the case of the Giampilieri catchment, Nat. Hazards Earth Syst. Sci., 12, 1295–1309, doi:10.5194/nhess-12-1295-2012, 2012.

Ardizzone, F., Basile, G., Cardinali, M., Casagli, N., Del Conte, S., Del Ventisette, C., Fiorucci, F., Garfagnoli, F., Gigli, G., Guzzetti, F., Iovine, G., Mondini, A. C., Moretti, S., Panebianco, M., Raspini, F., Reichenbach, P., Rossi, M., Tanteri, L., and Terranova, O.: Landslide inventory map for the Briga and the Giampilieri catchments, NE Sicily, Italy, J. Maps, 8, 176–180, doi:10.1080/17445647.2012.694271, 2012.

Atkinson, P., Jiskoot, H., Massari, R., and Murray, T.: Generalized linear modelling in geomorphology, Earth Surf. Proc. Land., 23, 1185–1195, 1998.

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

Bai, S.-B., Wang, J., Lü, G.-N., Zhou, P.-G., Hou, S.-S., and Xu, S.-N.: GIS-based and data-driven bivariate landslide-susceptibility mapping in the Three Gorges Area, China, Pedosphere, 19, 14–20, doi:10.1016/S1002-0160(08)60079-X, 2009.

Beguería, S.: Changes in land cover and shallow landslide activity: a case study in the Spanish Pyrenees, Geomorphology, 74, 196–206, doi:10.1016/j.geomorph.2005.07.018, 2006.

Brabb, E. E.: Innovative approaches to landslide hazard and risk mapping, in: Proceedings of the 4th International Symposium on Landslides, Toronto, Ontario, Canada, 16–21 September, Canadian Geotechnical Society, Toronto, Ontario, Canada, 1, 307–324, 1984.

Brenning, A.: 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, 2005.

Carrara, A., Cardinali, M., Guzzetti, F., and Reichenbach, P.: GIS technology in mapping landslide hazard, in: Geographical Information Systems in Assessing Natural Hazards, edited by: Carrara, A. and Guzzetti, F., Kluwer Academic Publisher, Dordrecht, the Netherlands, 135–175, 1995.

Carrara, A., Crosta, G., and Frattini, P.: Geomorphological and historical data in assessing landslide hazard, Earth Surf. Proc. Land., 28, 1125–1142, doi:10.1002/esp.545, 2003.

Carrara, A., Crosta, G., and Frattini, P.: Comparing models of debris-flow susceptibility in the alpine environment, Geomorphology, 94, 353–378, doi:10.1016/j.geomorph.2006.10.033, 2008.

Chung, C.-J. F. and Fabbri, A. G.: Validation of spatial prediction models for landslide hazard mapping, Nat. Hazards, 30, 451–472, doi:10.1023/B:NHAZ.0000007172.62651.2b, 2003.

Conoscenti, C., Di Maggio, C., and Rotigliano, E.: GIS analysis to assess landslide susceptibility in a fluvial basin of NW Sicily (Italy), Geomorphology, 94, 325–339, doi:10.1016/j.geomorph.2006.10.039, 2008a.

Conoscenti, C., Maggio, C., and Rotigliano, E.: Soil erosion susceptibility assessment and validation using a geostatistical multivariate approach: a test in Southern Sicily, Nat. Hazards, 46, 287–305, doi:10.1007/s11069-007-9188-0, 2008b.

Conrad, O.: SAGA – Entwurf, Funktionsumfang und Anwendung eines Systems für Automatisierte Geowissenschaftliche Analysen, Dissertation, University of Göttingen, 2007.

NHESSD

3, 1731–1774, 2015

Predicting storm triggered debris flow events

M. Cama et al.

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

Costanzo, D., Rotigliano, E., Irigaray, C., Jiménez-Perálvarez, J. D., and Chacón, J.: 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, 2012b.

Costanzo, D., Chacón, J., Conoscenti, C., Irigaray, C., and Rotigliano, E.: Forward logistic regression for earth-flow landslide susceptibility assessment in the Platani river basin (southern Sicily, Italy), Landslides, 11, 639–653, doi:10.1007/s10346-013-0415-3, 2014.

Del Ventisette, C., Garfagnoli, F., Ciampalini, A., Battistini, A., Gigli, G., Moretti, S., and Casagli, N.: An integrated approach to the study of catastrophic debris-flows: geological hazard and human influence, Nat. Hazards Earth Syst. Sci., 12, 2907–2922, doi:10.5194/nhess-12-2907-2012, 2012.

Frattini, P., Crosta, G., and Carrara, A.: Techniques for evaluating the performance of landslide susceptibility models, Eng. Geol., 111, 62–72, doi:10.1016/j.enggeo.2009.12.004, 2010.

Guzzetti, F., Reichenbach, P., Ardizzone, F., Cardinali, M., and Galli, M.: Estimating the quality of landslide susceptibility models, Geomorphology, 81, 166–184, doi:10.1016/j.geomorph.2006.04.007, 2006.

Guzzetti, F., Mondini, A. C., Cardinali, M., Fiorucci, F., Santangelo, M., and Chang, K. T.: Landslide inventory maps: new tools for an old problem, Earth-Science Rev., 112, 42–66, 2012.

Heckmann, T., Gegg, K., Gegg, A., and Becht, M.: 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, 2014.

Hosmer, D. W. and Lemeshow, S.: Applied Logistic Regression, 2nd Edn., Wiley, New York, doi:10.1198/tech.2002.s650, 2000.

Hungr, O., Evans, S. G., Bovis, M. J., and Hutchinson, J. N.: A review of the classification of landslides of the flow type, Environ. Eng. Geosci., 7, 221–238, doi:10.2113/gseegeosci.7.3.221, 2001.

Hungr, O., Leroueil, S., and Picarelli, L.: The Varnes classification of landslide types, an update, Landslides, 11, 167–194, doi:10.1007/s10346-013-0436-y, 2014.

Hutchinson, J. N.: General Report: Morphological and geotechnical parameters of landslides in relation to geology and hydrogeology, Proceedings, in: Fifth International Symposium on Landslides, edited by: Bonnard, C., 1, Balkema, Rotterdam, 3–35, 1988.
Lentini, F., Carbone, S., and Messina, A.: Carta Geologica d'Italia scala 1 : 50.000 Foglio 601 "Messina-Reggio di Calabria", con note illustrative. Apat Agenzia per la protezione dell'ambiente e per i servizi tecnici, Dipartimento Difesa del Suolo – Servizio geologico d'Italia, Firenze, 2007.
Lombardo, L., Cama, M., Maerker, M., and Rotigliano, E.: A test of transferability for landslides susceptibility models under extreme climatic events: application to the Messina 2009 disaster, Nat. Hazards, 74, 1951–1989, doi:10.1007/s11069-014-1285-2, 2014.
Menard, S.: Applied Logistic Regression Analysis, SAGE,Thousand Oaks, California, 1–120, doi:10.4135/9781412983433, 2002.
Meusburger, K. and Alewell, C.: On the influence of temporal change on the validity of landslide susceptibility maps, Nat. Hazards Earth Syst. Sci., 9, 1495–1507, doi:10.5194/nhess-9-1495-2009, 2009.
Messina, A., Somma, R., Macaione, E., Carbone, G., and Careri, G.: Peloritani continental crust composition (southern Italy): geological and petrochemical evidences, Boll. Soc. Geol. Ital., 123, 405–444, 2004.
Moore, I. D., Gessler, P. E., Nielsen, G. A., and Peterson, G. A.: Soil attribute prediction using terrain analysis, Soil Sci. Soc. Am. J., 57, 443–452, doi:10.2136/sssaj1993.572NPb, 1993
Nefeslioglu, H. A., Gokceoglu, C., and Sonmez, H.: 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, 2008.
Ohlmacher, G. C.: Plan curvature and landslide probability in regions dominated by earth flows and earth slides, Eng. Geol., 91, 117–134, doi:10.1016/j.enggeo.2007.01.005, 2007.
Ohlmacher, G. C. and Davis, J. C.: 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, 2003.
Rakotomalala, R.: TANAGRA: un logiciel gratuit pour l'enseignement et la recherche, in: 5émes Journées d'Extraction et Gestion des Connaissances proceedings, EGC-2005, 697–702, Paris, 2005

Reichenbach, P., Busca, C., Mondini, A. C., and Rossi, M.,: The influence of land use change on landslide susceptibility zonation: the Briga Catchment Test Site (Messina, Italy), Environ. Manage., 54, 1372–1384, doi:10.1007/s00267-014-0357-0, 2014.

Rossi, M., Guzzetti, F., Reichenbach, P., Mondini, A. C., and Peruccacci, S.: Optimal landslide susceptibility zonation based on multiple forecasts, Geomorphology, 114, 129–142, doi:10.1016/j.geomorph.2009.06.020, 2010.

Rotigliano, E., Agnesi, V., Cappadonia, C., and Conoscenti, C.: 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, 2011.

Süzen, M. L. and Doyuran, V.: A comparison of the GIS based landslide susceptibility assessment methods: multivariate versus bivariate, Environ. Geol., 45, 665–679, doi:10.1007/s00254-003-0917-8, 2004.

Van Den Eeckhaut, M., Reichenbach, P., Guzzetti, F., Rossi, M., and Poesen, J.: Combined landslide inventory and susceptibility assessment based on different mapping units: an example from the Flemish Ardennes, Belgium, Nat. Hazards Earth Syst. Sci., 9, 507–521, doi:10.5194/nhess-9-507-2009, 2009.

Varnes, D. J.: Slope movement: types and process, in: Landslides: Analyses and Control, edited by: Schuster, R. L. and Krizek, R. J., Sp. Rep., Transportation Research Board, Natural Academy of Science, Washington, 176, 11–33, 1978.

Von Ruette, J., Papritz, A., Lehmann, P., Rickli, C., and Or, D.: 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, 2011.

NHESSD

3, 1731–1774, 2015

Predicting storm triggered debris flow events

M. Cama et al.

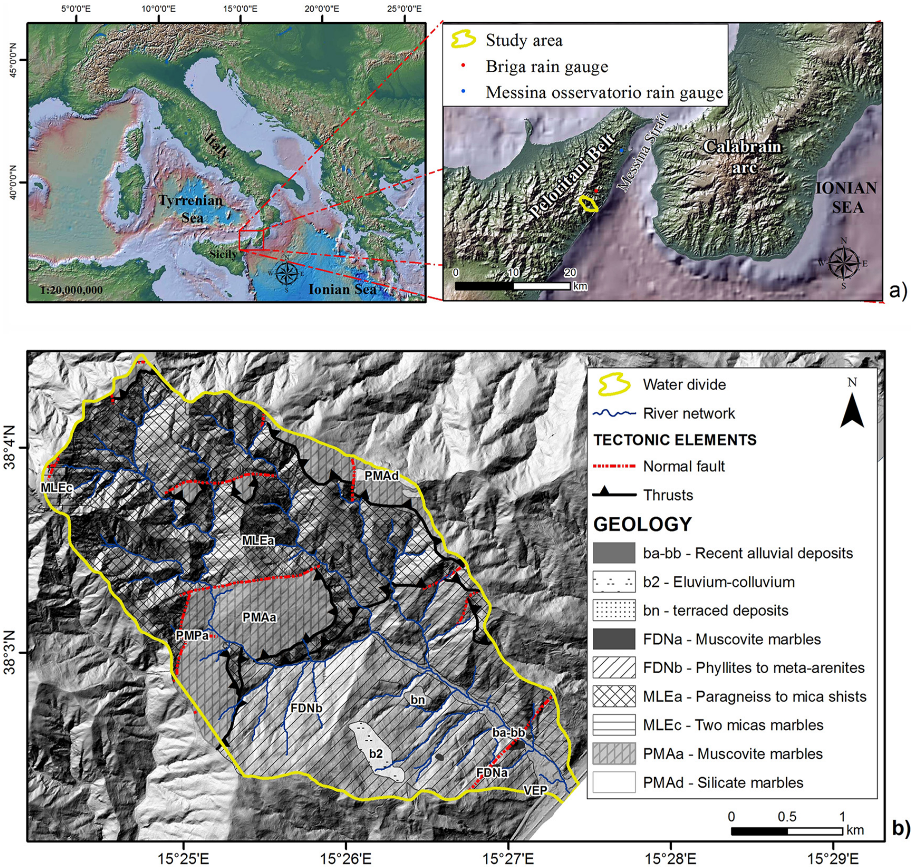

Figure 1. Setting of the study area.

NHESSD
3, 1731–1774, 2015

Predicting storm triggered debris flow events

M. Cama et al.

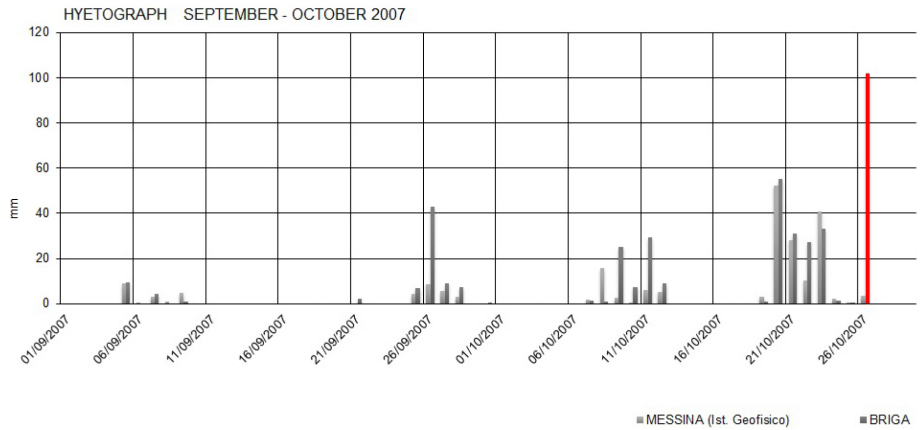

Figure 2. Hyetograph showing the daily precipitation during the month of October 2007. The 2007 event is clearly distinguishable on the 16 October registration.

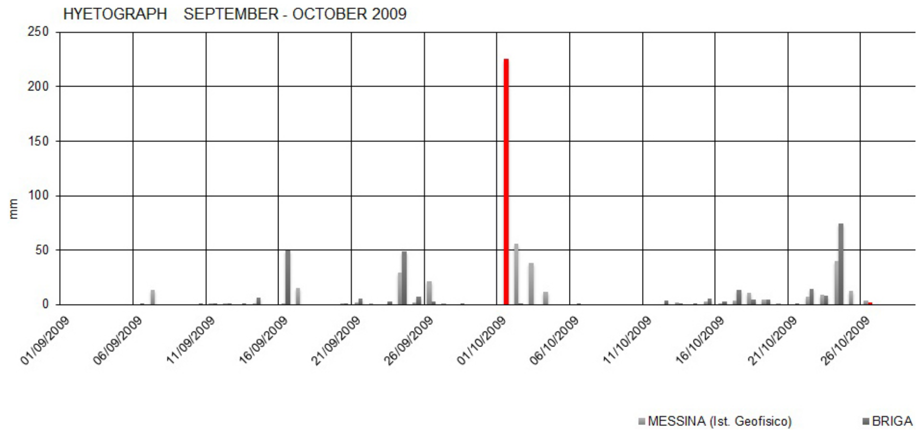

Figure 3. Hyetograph showing the daily precipitation during the months of September and October 2009.

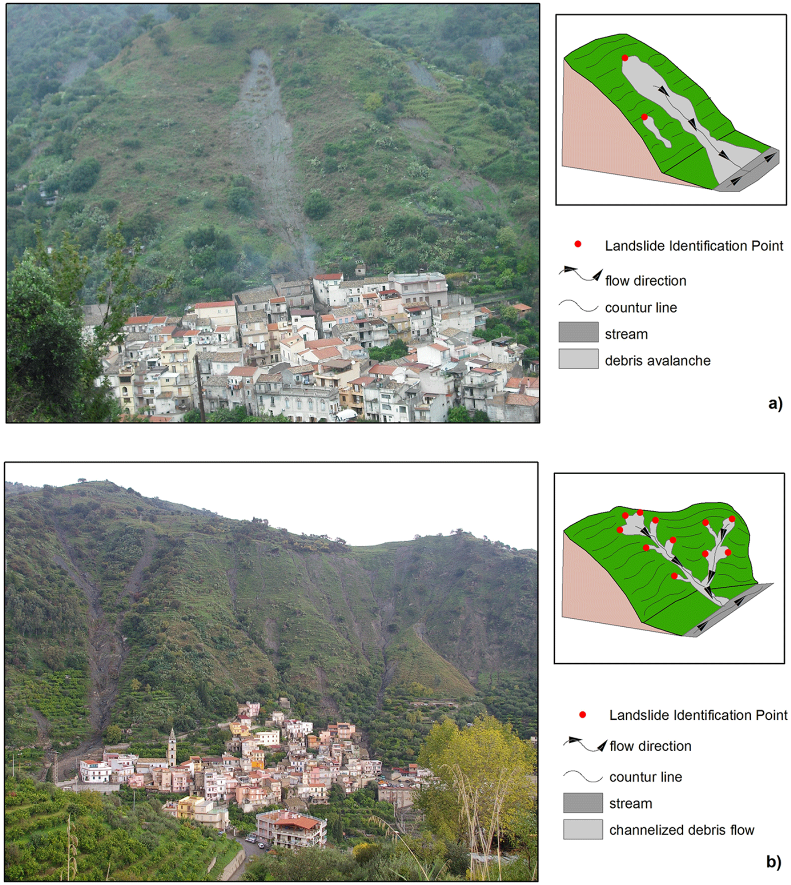

Figure 4. Overview of the area hit by the 2009 event: **(a)** Guidomandri village: debris avalanches are observable on the triangular facets parallel to the coast; **(b)** Itala village: channelized debris flows crossing the urbanized area.

NHESSD
3, 1731–1774, 2015

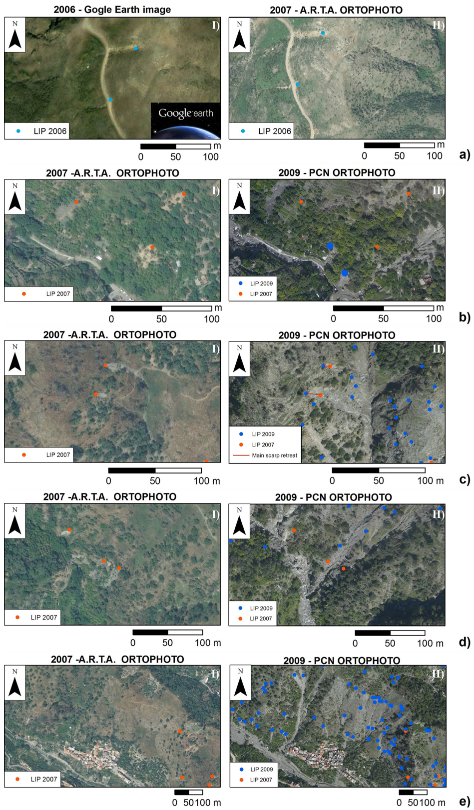

Figure 5. Comparison of morphologies between two different images resulting in five different cases: **(a)** shallow landslides recognized on the 2007 orthophoto but activated before the 2007 event; **(b)** shallow landslides activated in 2007 which did not reactivated or retreated in 2009; **(c)** shallow landslides which activated in 2007 that retreated or reactivated in 2009; **(d)** shallow landslides activated in 2007 which have been completely included in 2009; **(e)** shallow landslides activated in 2009.

NHESSD
3, 1731–1774, 2015

Predicting storm triggered debris flow events

M. Cama et al.

Title Page
Abstract Introduction
Conclusions References
Tables Figures
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

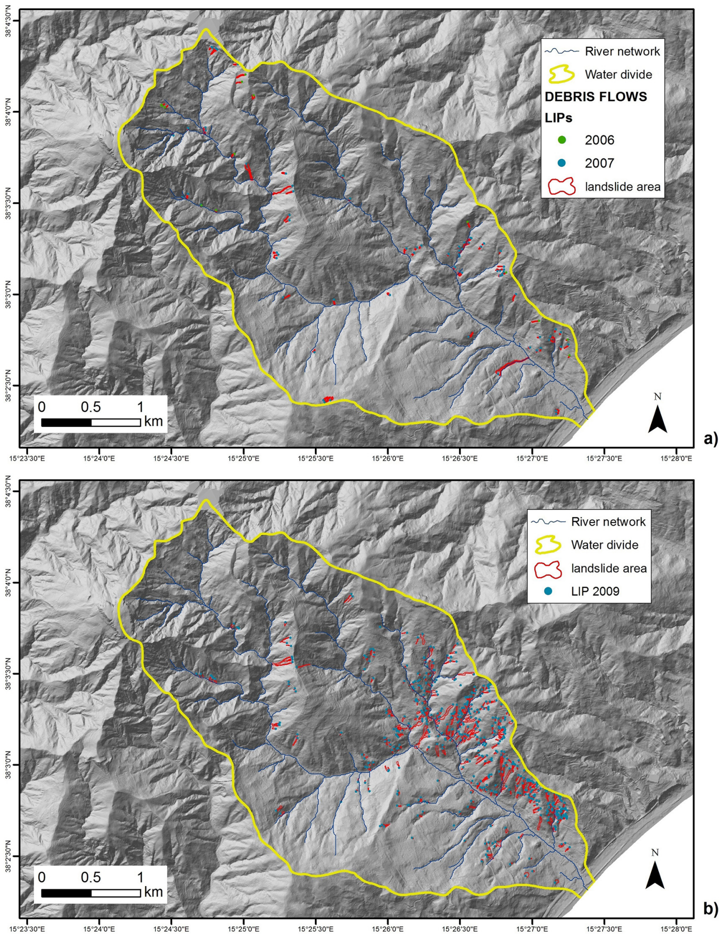

Figure 6. Debris flow event inventories: **(a)** 2007 inventory containing 73 phenomena; **(b)** 2009 event inventory containing 616 phenomena.

NHESSD
3, 1731–1774, 2015

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◄| |►
◄ ►
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

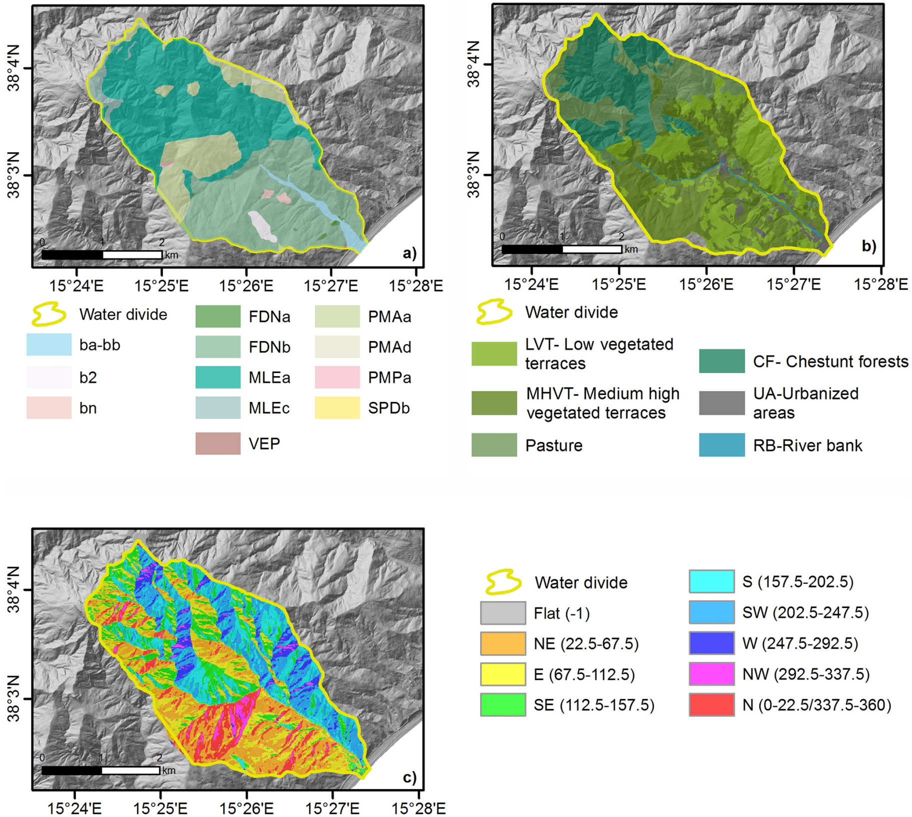

Figure 7. Discrete variables: **(a)** outcropping lithology (GEO; see Fig. 1 for description); **(b)** land use (USE); **(c)** aspect (ASP).

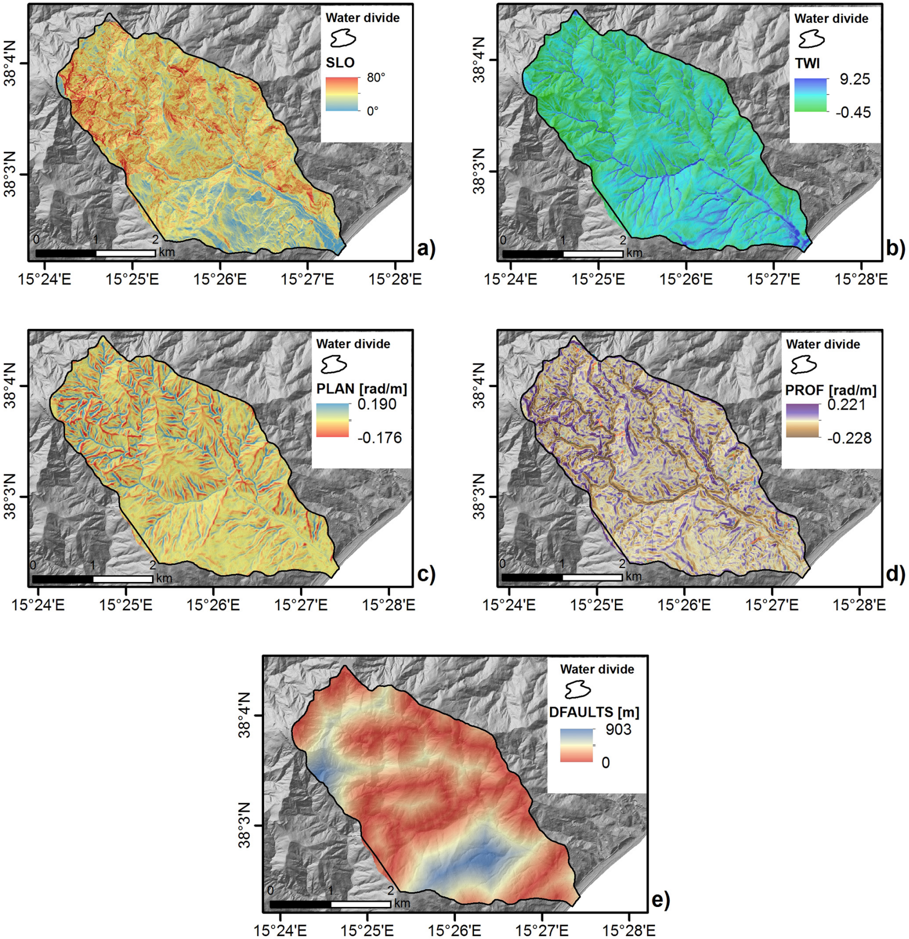

Figure 8. Continuous variables: **(a)** slope; **(b)** opographic wetness index; **(c)** plan curvature; **(d)** profile curvature; distance from tectonic elements.

NHESSD
3, 1731–1774, 2015

0.5 cut-off model accuracy (10 x 10 cross validation)

2007 self-validating 90/10% models			Error rate			
			mean	0.33	dev. Std	0.042
Values prediction			Confusion matrix			
Value	Recall	1-Precision		NO	YES	Sum
NO	0.653	0.329	NO	136.7	72.7	209.4
YES	0.680	0.336	YES	67.3	143.3	210.6
			Sum	204	216	420

2009 self-validating 90/10% models			Error rate			
			mean	0.23	dev. Std	0.017
Values prediction			Confusion matrix			
Value	Recall	1-Precision		NO	YES	Sum
NO	0.774	0.229	NO	1428.6	416.4	1845
YES	0.770	0.227	YES	425	1420	1845
			Sum	1853.6	1836.4	3690

ROC-AUCs (self-validating 90/10% models)

		1	2	3	4	5	6	7	8	9	10	MEAN	ST.DEV.
2007	Model fitting	0.85	0.85	0.80	0.80	0.81	0.85	0.83	0.80	0.78	0.80	0.82	0.027
	Prediction skill	0.81	0.52	0.97	0.81	0.87	0.70	0.81	0.90	0.65	0.85	0.79	0.131
2009	Model fitting	0.86	0.87	0.89	0.87	0.88	0.88	0.88	0.87	0.87	0.88	0.87	0.009
	Prediction skill	0.85	0.90	0.87	0.86	0.85	0.86	0.88	0.85	0.86	0.86	0.86	0.015

Figure 9. Error rate and AUC for self-validating (cross validation) 2007 and 2009 models.

NHESSD
3, 1731–1774, 2015

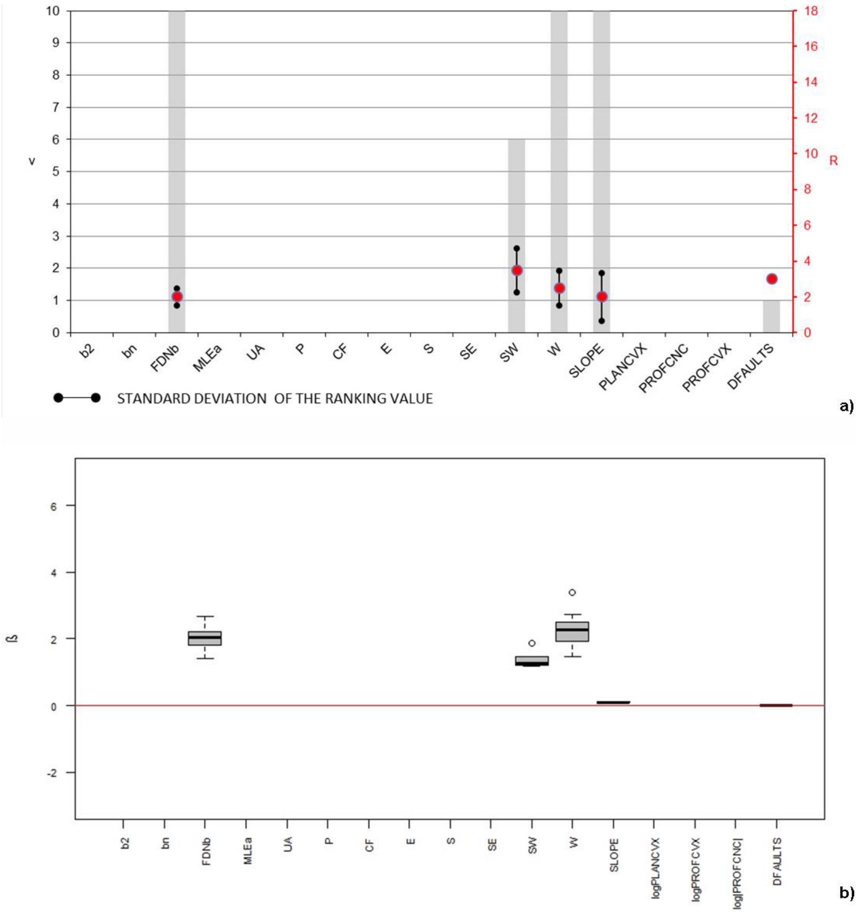

Figure 10. Selected variables for the 2007 suites of models: **(a)** ranking and frequency; **(b)** β values (the β values of the curvatures is expressed in logarithm).

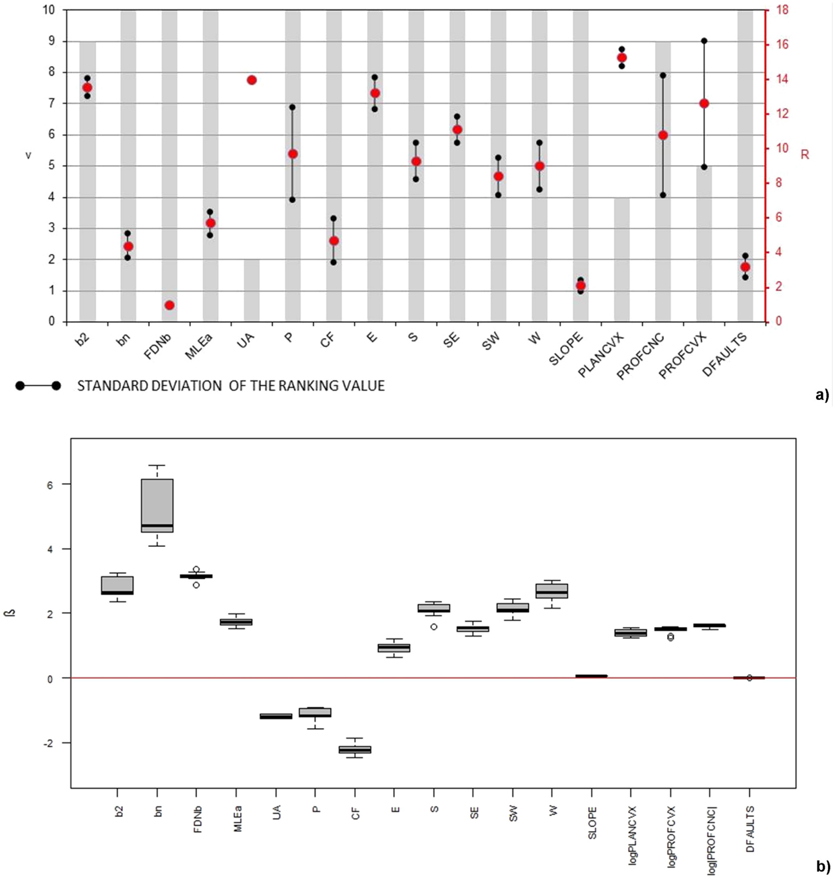

Figure 11. Selected variables for the 2009 suites of models: **(a)** ranking and frequency; **(b)** β values.

NHESSD
3, 1731–1774, 2015

Predicting storm triggered debris flow events

M. Cama et al.

Title Page
Abstract Introduction
Conclusions References
Tables Figures
◀◀ ▶▶
◀ ▶
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

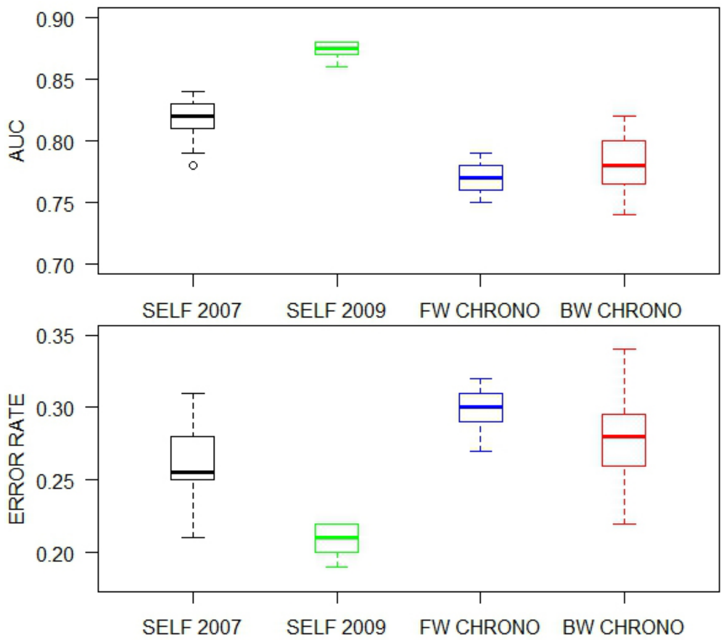

Figure 12. Distribution of the AUC and error rate values calculated on the 10 replicates for 2007 and 2009 modelling and 100 models during the chrono-validation process.

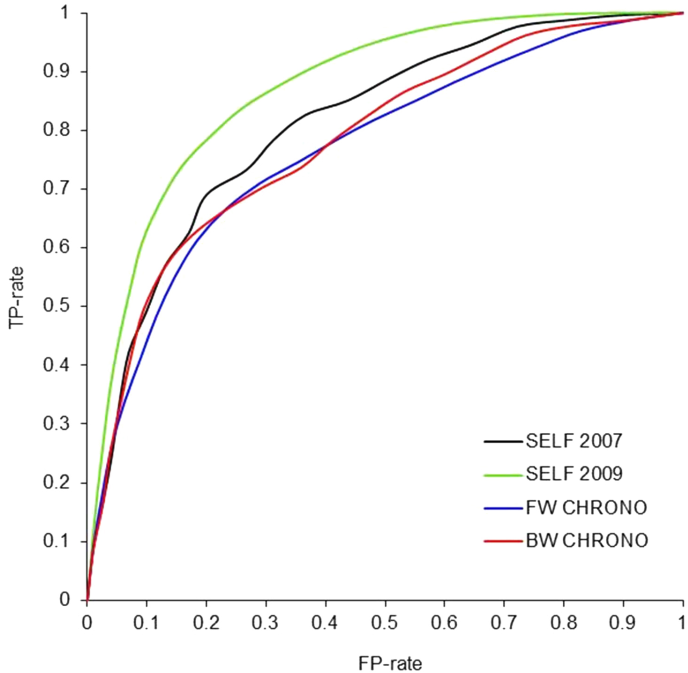

Figure 13. Mean ROC curves.

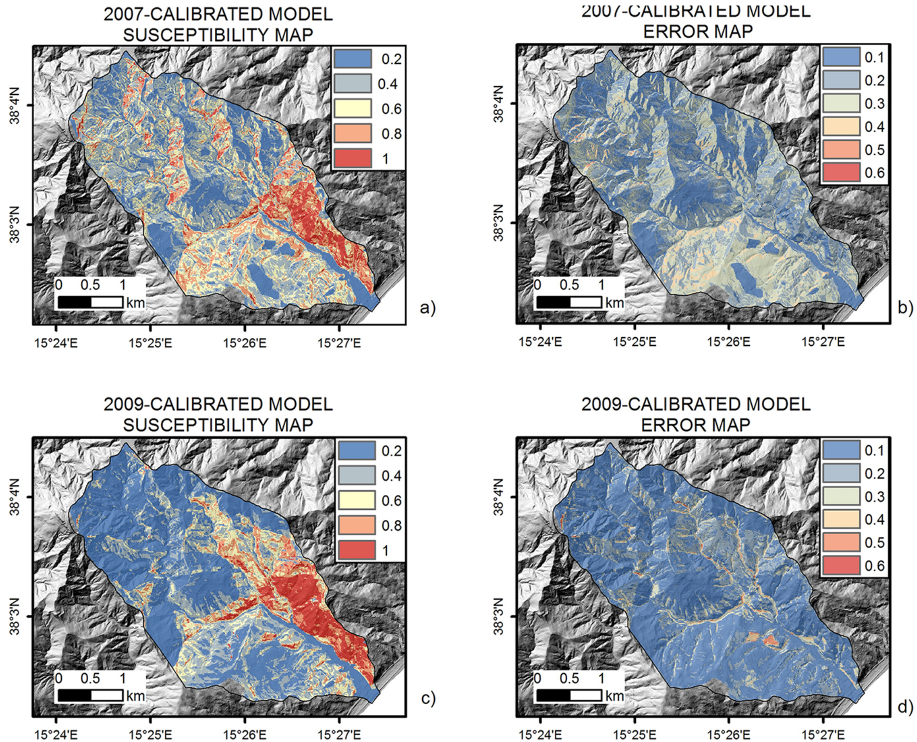

Figure 14. Susceptibility and error maps for the 2007- and the 2009-calibrated models: **(a, c)** mean susceptibility; **(b, d)** error maps.

NHESSD
3, 1731–1774, 2015

Title Page
Abstract Introduction
Conclusions References
Tables Figures
◂I I▸
◂ ▸
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

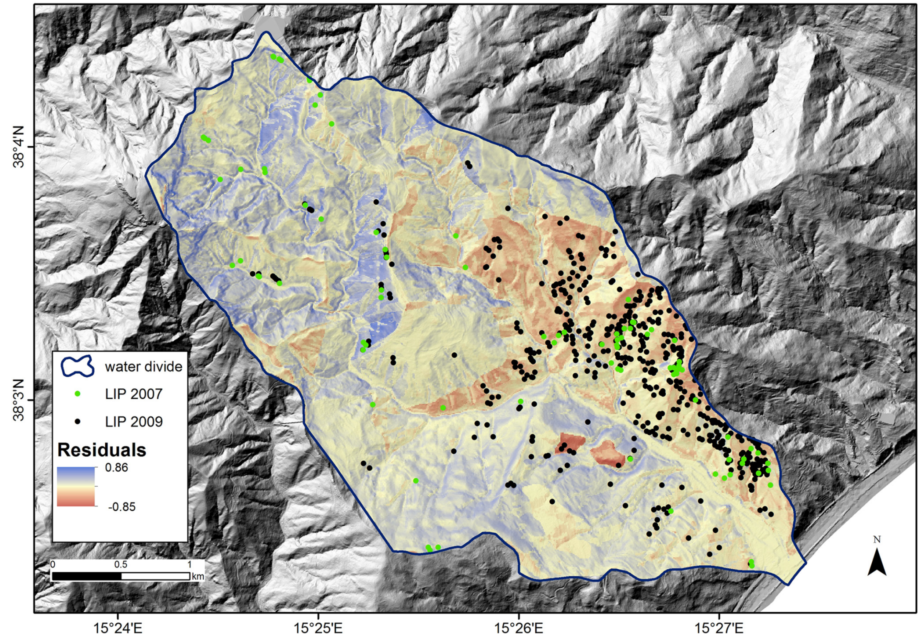

Figure 15. Map of residuals calculated as percentage differences between the two (2007 and 2009) mean susceptibilities.

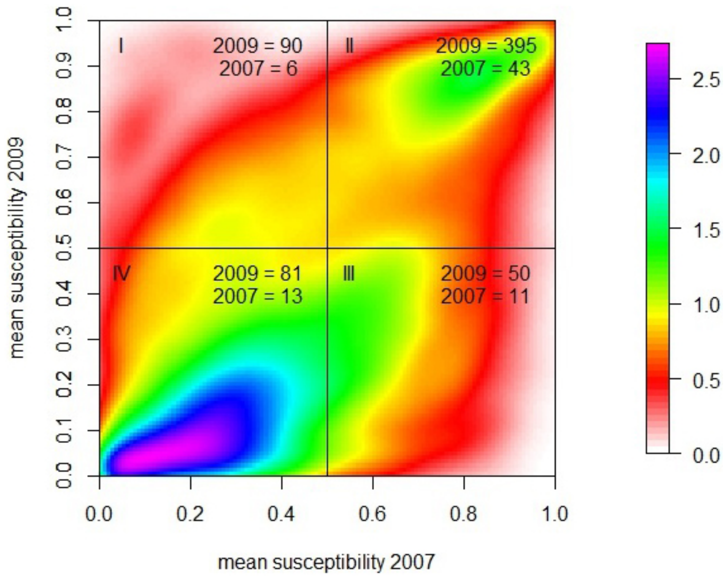

Figure 16. Dispersion density plot calculated using 2-D Binned Kernel Density algorithm (*xy* range for density calculation 0.045). Positive cases for 0.5 cut-off values are reported for the two inventory events.