Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-457
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 24 January 2018

Natural Hazards and Earth System Sciences
Discussions

Estimation of the susceptibility of a road network to shallow landslides with the integration of the sediment connectivity

Massimiliano Bordoni[1], M. Giuseppina Persichillo[1], Claudia Meisina[1], Stefano Crema[2], Marco Cavalli[2], Carlotta Bartelletti[3], Yuri Galanti[3], Michele Barsanti[4], Roberto Giannecchini[3], Giacomo D'Amato Avanzi[3]

[1]Department of Earth and Environmental Sciences, University of Pavia, Pavia, 27100, Italy
[2]Research Institute for Geo-Hydrological Protection, National Research Council, Padova, 35127, Italy
[3]Department of Earth Sciences, University of Pisa, Pisa, 56126, Italy
[4]Department of Civil and Industrial Engineering, University of Pisa, Pisa, 56126, Italy

Correspondence to: Massimiliano Bordoni (massimiliano.bordoni01@universitadipavia.it)

Abstract. Landslides causes severe damages to the road network of a hit zone, in terms of both direct (partial or complete destruction of a road trait, blockages) and indirect (traffic restriction, cut-off of a certain area) costs. Thus, the identification of the parts of the road network which are more susceptible to landslides is fundamental to reduce the risk to the population potentially exposed and the money expense caused by road damaging. For these reasons, this paper aimed to develop and test a data-driven model based on the Genetic Algorithm Method for the identification of road sectors that are susceptible to be hit by shallow landslides triggered in slopes upstream to the infrastructure. This work also analyzed the importance of considering or not the sediment connectivity on the estimation of the susceptibility. The study was carried out in a catchment of north-eastern Oltrepò Pavese (northern Italy), where several shallow landslides affected roads in the last 8 years. The random partition of the dataset used for building the model in two parts (training and test subsets), within a 100-fold bootstrap procedure, allowed to select the most significant explanatory variables, providing a better description of the occurrence and distribution of the road sectors potentially susceptible to damages induced by shallow landslides. The presented methodology allows the identification, in a robust and reliable way, of the most susceptible road sectors that could be hit by sediments delivered by landslides. The best predictive capability was obtained using a model which took into account also the index of connectivity, calculated according to a linear relationship. Most susceptible road traits resulted to be located below steep slopes with a limited height (lower than 50 m), where sediment connectivity is high. Different scenarios of land use were implemented in order to estimate possible changes in road susceptibility. Land use classes of the study area were characterized by similar connectivity features with a consequent loss of variations also on the susceptibility of the road networks according to different scenarios of distribution of land cover. Larger effects on sediment connectivity and, as a consequence on road susceptibility, could be due to modifications in the morphology of the slopes (e.g. drainage system, modification of the slope angle) caused by the abandonment or by the recovery of cultivations. The results of this research demonstrate the ability of the developed methodology in the assessment of susceptible roads. This could give to the managers of an infrastructure information on the

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-457
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 24 January 2018

Natural Hazards and Earth System Sciences Discussions

criticality of the different road traits, thereby allowing attention and economic budgets to be shifted towards the most critical assets, where structural and non-structural mitigation measures could be implemented.

Keywords: roads, shallow landslides, susceptibility

1 Introduction

Landslides are an important geohazard in many regions of the world, causing severe economic damages each year in the order of hundreds of billions dollars (Zezere et al., 2007; Salvati et al., 2014; Gariano and Guzzetti, 2016). Slope instability induces significant damage, deaths and economic losses to infrastructures, in particular to roads (Van Westen et al., 2006; Klose et al., 2015). The main negative consequences of instability phenomena on roads are (Bil et al., 2014): i) the partial or complete destruction of a road trait, which can also provoke human losses; ii) the traffic restriction due to the blockage of a hit road, which may affect the entire network causing congestion; iii) the cut-off of certain areas that cannot be reached by alternative routes.

Thus, it is fundamental to identify which sectors of a road network are more susceptible to landslides, in order to reduce the risk to the population potentially exposed and the money expense caused by road damaging. This aim is particularly important, also because several researches (Nemry and Demirel, 2012; Matulla et al., 2017) stressed that the exposure of road networks to slope instabilities could increase as a consequence of the climate change and of the economic rising income in different countries.

According to the geomorphological and triggering features, landslides affecting roads can be distinguished in: i) landslides in correspondence of the infrastructure; ii) landslides triggered in a natural or an engineered hillslope upstream to the road, whose transportation and/or accumulation zone hit the infrastructure.

The triggering mechanisms of the first landslide type are strictly related to local hydrological and geotechnical settings that are related to the road presence. These factors generally highlight an incorrect construction or management of the infrastructure regardless of the natural features of the slopes where the infrastructure was built (Sidle and Ochiai, 2006; Muenchow et al., 2012; D'Amato Avanzi et al., 2013; Brenning et al., 2015). On the other hand, the triggering mechanism and landslide runout of the second landslide type can be related to the geological, geomorphological and hydrological predisposing factors of the natural or engineered slopes upstream to the roads. Furthermore, these events are the most widespread in terms of affected routes, causing in many cases the involvement of extended sectors of hilly and mountainous road networks (Quinn et al., 2010; Bil et al., 2014).

In recent years, several data-driven methodologies were built to identify the susceptible sectors of a road network towards landslides (Budetta, 2004; Jaiswal et al., 2010a, 2010b, 2011; Quinn et al., 2010; Michoud et al., 2012; Bil et al., 2014; Ramesh and Anbazhagan, 2015; Pellicani et al., 2017). These methods are based on quantitative statistical relationships between predisposing factors and a response variable, assuming that an event is most likely to occur under similar ground conditions

to previous events (Varnes, 1984). They present the advantage of being more objective and easily applicable at different scales (from site-specific to regional), as well as capable of managing large sets of predisposing factors (Corominas et al., 2014).

Data-driven models used for the characterization of susceptible routes were based on a multivariate analysis (Dai and Lee 2002; Chen and Wang 2007), which predicts the spatial distribution of roads hit by landslides through the estimation of the relations and the relative weight between the predisposing factors and the response variable (roads affected by landslides). Such methods do not consider the non-linear relations between the predisposing factors and the response variables. However, the non-linearity of the system should be considered, since the changing in the environmental and geological conditions leads to a consequent interaction of the mobilized materials with roads (Goetz et al., 2011). Moreover, neglecting a possible non-linearity in the model could decrease its predictive performance, due to a limitation in highlighting the complex behaviors of the phenomena (Phillips 2003, 2006). Thus, it could be useful implementing a model based on a non-linear regression technique, such as the Generalized Additive Model (GAM; Hastie and Tibshirani, 1990).

Furthermore, the methods previously developed did not take into account for the potential slope sediments mobilized by the landslide triggering, to reach the road network in downstream area. This aspect is instead well described by the amount in sediment connectivity, which influences the path and the travel distance of the materials mobilized by a slope failure till a potential natural or anthropogenic barrier (e.g. a river, a road) (Cavalli et al., 2013; Tarolli and Sofia, 2016; Persichillo et al., 2018). In this way, the landslide runout can be estimated and inserted in the modeling of roads susceptibility, without employing numerical or physically-based methods which require several rheological and geotechnical data not easily measurable for the slope materials (Hungr, 1995; Fannin and Wise, 2001; Pastor et al., 2014; Fan et al., 2017).

It is also worth noting that scenarios of road susceptibility distributionrelated to the modifications of land use in a particular area were not considered so far. However, land use changes can have significant impacts both on the locations of landslides triggering zones (Glade, 2003; Begueria, 2006; Reichenbach et al., 2014; Persichillo et al., 2017) and on the connectivity of the mobilized sediments (Foerster et al., 2014; Lopez-Vicente et al., 2013, 2016). Thus, susceptibility scenarios of different land use distributions may allow to identify land management practices able to reduce the slope instability which can induce damages to roads.

For these reasons, a non-linear data-driven method, for the identification of road network sectors more susceptible to shallow landslides triggered in slopes upstream to the infrastructure, was developed and tested. The main objectives of the paper are: i) the development and the test of a data-driven non-linear methodology, based on the GAM, able to identify the relations between predisposing and response variables for the assessment of the road sectors susceptible to shallow landslides triggered in slope upstream to the infrastructure; ii) the evaluation of the importance of considering or neglecting sediment connectivity in the susceptible road segments modelling; iii) the analysis of the effects resulting from different scenarios of land use distribution on the routes distribution that could be affected by shallow landslides.

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-457
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 24 January 2018

Natural Hazards and Earth System Sciences Discussions

2 The study area

The analysis was carried out in a catchment located between Scuropasso river and Versa river catchments, in Oltrepò Pavese, in the northern termination part of the Italian Apennines (Fig. 1). The study area is 14 km^2 wide and presents an elevation range between 88 and 295 m a.s.l. The morphological structure is typical of the Pede-Apennine margin of Oltrepò Pavese and it is closely related to both the lithology and the tectonic/neotectonic setting of the Apennine margin. It is characterized by a medium-high slope gradient, with slope angles higher than 10°, with prominent altimetric irregularities along ridge lines and channel network in narrow valleys (Bordoni et al., 2015). Bedrock is characterized by a Mio-Pliocenic succession formed by medium low-permeable arenaceous conglomeratic materials (Monte Arzolo Sandstones, Rocca Ticozzi Conglomerates) overlying impermeable silty-sandy marly bedrock (Montù Beccaria Formation, Sant'Agata Fossili Marls) and evaporitic chalky marls and gypsum (Gessoso-Solfifera Formation) (Vercesi and Scagni, 1984). Superficial soils, derived by bedrock weathering, are mainly clayey-sandy silts and clayey-silty sands. Soil depth ranges from a few centimeters to values lower than 2.5 m.

The study area is characterized by a traditional viticulture vocation, with grapevine cultivation that represents the main economic branch of this zone. Till the 1980s, more than 90% of the territory was cultivated with vineyards, where manual cultivation practices predominated (Fig. 2a, b, c). This situation represented the highest diffusion of vineyards in the study area, identifying all the hillslopes that are effectively adapt for grapevine implantation and cultivation. Instead, in the last 40 years, more than 40% of previously cultivated slopes were abandoned, with a correspondent progressive increase in woodlands (+13% from 1980 to 2007-2015) and in uncultivated areas generally composed by shrubs and grasses (+10% from 1980 to 2007-2015) (Fig. 2a, d). In the period between 2007-2015, land use changes significantly decreased and land use classes distribution kept steady. In 2007-2015 time span, 49% of the area was occupied by vineyards, 10% by uncultivated areas, 16% by woodlands and 16% by urban areas (Fig. 2d). Other land use classes are present in a percentage lower than 5%.

This abandonment was due to the conversion from manual to mechanical cultivation practices, that increased the difficulties in the maintenance of vineyards, especially for those located on very steep slopes (> 25°) (Persichillo et al., 2017).

These land use changes reflected the efforts made to achieve a greater efficiency through mechanization of agricultural practices for improving economic productivity. Moreover, societal changes, together with the decreasing number of people actively cultivating the area, caused a reduction in land care practices and maintenance works in both abandoned and still cultivated vineyards (Persichillo et al., 2017, 2018).

A primary road network (81 km long and generally 3–5 m wide) crosses the study area; it is composed of provincial and municipal roads that connect different villages and towns (Fig. 3). The road sectors were built in correspondence of the valley floors or hillside, cutting a portion of a hillslope in correspondence of its medium part realizing a halfway road. Furthermore, in the case of halfway roads, a 3–5 m height trench was built upstream to the road sector.

This area was recently affected by several shallow landslides triggered by intense rainfall events (Bordoni et al., 2015). The most important one occurred in 27–28 April 2009 (160 mm/62 h), and induced 532 failures. Other shallow landslides occurred

during the events of March/April 2013 and of 28 February–2 March 2014, triggering 19 and 18 shallow failures, respectively. Lower numbers of phenomena reflect the lower amount of rainfall recorded during these events (40 mm in about 30-50 h in March/April 2013 events; 69 mm in 42 h in 28 February–2 March 2014 event).

These landslides had an average length of about 35 m and their area varied from a minimum of 13 m^2 to a maximum of almost 9,000 m^2, with an average of about 477 m^2. The failure surface was mainly detected between 0.9 and 1 m from the ground level, generally in correspondence between the soil-bedrock contact. 30% of these shallow landslides triggered in vineyards, while an equal percentage of phenomena developed in woodlands or uncultivated areas. According to Cruden and Varnes (1996) classification, most of the shallow landslides can be classified as roto-translational slides evolved into flows, with width/length ratio > 1. Moreover, 24 failures (5% of the total number), were roto-translational slides affecting the trench in correspondence of a cut of a halfway road (B2 type; Zizioli et al., 2013; Persichillo et al., 2018).

Besides the partial or complete destruction of cultivated vineyards, the landslides significantly affected the road network with severe damage (Fig. 3) regarding the partial or complete destruction of road traits, debris accumulation and blockage, causing traffic restriction and the cut-off of villages and towns.

In particular, 2.5 km of the principal road network was affected by shallow landslides in the last years. 134 shallow landslides (23%) hit roads: 24 of them (15%) were roto-translational slides developed in correspondence of the trench upstream the road trait, while the remaining 90 phenomena (85%) were shallow landslides triggered in slopes upstream the routes on cultivated or abandoned hillslopes. The length of the road sectors hit by a shallow landslide ranged between 2 and 94 m.

3 Methods

3.1 Development and test of the data-driven model

3.1.1 Predictor variables

A data-driven methodology based on GAM was implemented for the assessment of roads that could be hit by shallow landslides. A schematic flow-chart of this methodology is shown in Fig. 4.Such procedure is similar to one proposed by Persichillo et al. (2016) for the assessment of the shallow landslide susceptibility in different settings. It was refined in this paper for the application to roads susceptibility towards landslides. In particular, different predictor variables and response variables were considered, according to their influence on the possible interaction between landslide mobilized materials and the road network located downstream.

In the model, 11 predictor variables were identified. 8 of these parameters were extracted by a 1 m resolution LiDAR-derived Digital Elevation Model (DEM), through SAGA GIS (System for Automated Geoscientific Analyses). The DEM was available from the Italian Ministry of Environment and Protection of the Land and Sea, following the realization of the Piano Straordinario di Telerilevamento Ambientale (Extraordinary Plan of Environmental Remote Sensing - PST-A). These attributes were: slope angle (SL), slope aspect (ASP), slope curvature (CURV), slope length (LEN), slope height (HEI), catchment area (CA), catchment slope (CS) and topographic wetness index (TWI).

SL, ASP, and CURV were calculated through Zevenbergen and Thorne (1987)' approximations. SL strongly controls the velocity of the material mobilized by a shallow landslide, thus its capacity of traveling for long distances from the source areas (Fannin and Wise, 2001; Catani et al., 2013; Fathani et al., 2017). ASP influences the soil moisture and the vegetation growth, that can have a key role on the susceptibility of a slope to shallow failures (Van Westen et al., 2008; Jaiswal et al., 2010a). CURV influences the amount of water runoff, the rate of underground water movement and the potential rates of sedimentation and erosion (Dai et al., 2002; Kritikos and Davis, 2015).

LEN and HEI are key parameters for the estimation of the distance travelled by a landslide from its source area and of the velocity of the displaced material when it hits an infrastructure (Bathurst et al., 1997; Chau et al., 2004; Martinovic et al., 2016).

Multiple-flow direction algorithm (Quinn et al., 1991) was used to obtain CA and CS. CA is used as a proxy for soil moisture and soil depth, thus for the potential amount of materials that can be mobilized by the shallow landslide and that can reach an infrastructure (Brenning et al., 2015). CS conditions the destabilizing forces upstream that can provoke the development of a landslide (Brenning et al., 2015; Persichillo et al., 2016). TWI highlights the water fluxes along the slopes and the position of the accumulation points in a catchment (Seibert et al., 2007).

Along with the DEM-derived predictor variables, the Euclidean distance from shallow landslide source area (DIST) was calculated, considering the lowest distance between the landslide source area and a considered road trait. This parameter is adapted for slopes with homogeneous gradient, aspect and curvature and it is important to understand the capacity of the mobilized materials to travel along a slope and to reach a route located downstream (Bil et al., 2014; Brenning et al., 2015). The source area of each slope failure was extracted through the procedure of Galve et al. (2015), selecting 25% of the landslide area in correspondence to the highest elevations.

Bedrock geology (GEO) was also considered as predictor. GEO influences the geomechanical, geotechnical, rheological and hydrological properties of the soil, which have effects on the runout of a landslide (Hungr, 1995; Pastor et al., 2014). GEO was obtained from the geological map of the studied catchment, which was obtained by the Department of Earth and Environmental Sciences of University of Pavia through field surveys.

Different authors (Budetta, 2004; Jaiswal et al., 2010a, 2010b, 2011; Quinn et al., 2010; Michoud et al., 2012; Bil et al., 2014; Ramesh and Anbazhagan, 2015; Pellicani et al., 2017) had already used some of the previously described predictor variables in different data-driven model aiming to assess roads susceptible to be hit by shallow landslides. Until now, sediment connectivity has not been considered yet as a predictor variable influencing the susceptibility of road network. Persichillo et al. (2018) demonstrated that, in two catchments of Oltrepò Pavese, the road sectors hit by the materials mobilized by shallow landslides occurred upstream are the ones located close to slopes characterized by the lowest or the highest values of sediments connectivity along the entire catchment. In order to verify the potential influence of this parameter in discriminating the susceptible road sectors, an index of sediment connectivity within the predictor variables of the model was inserted.

The index of sediment connectivity (IC), defined by Borselli et al. (2008), evaluates the potential connection between hillslopes and features, which act as targets or storage areas (sinks) for mobilized sediments (e.g., channels, basin outlet, lakes, road

network). In the proposed model, IC, calculated according to the approach by Cavalli et al. (2013), was implemented for a better characterization of surface processes and properties and to exploit a high-resolution DEM. For further details on the changes introduced in the IC calculation following this scheme, we refer to Cavalli et al. (2013) and Crema and Cavalli (2018).

IC is calculated according to Eq. (1) combining the upslope (D_{up}) and downslope (D_{dn}) components of connectivity, respectively:

$$IC = log_{10}\frac{D_{up}}{D_{dn}} \quad (1)$$

IC can have values in the range of $[-\infty, +\infty]$, with connectivity increasing for larger IC values (Cavalli et al., 2013). IC was calculated through the stand-alone application SedInConnect 2.3 (Crema and Cavalli, 2018).

In the calculation of IC, both the D_{up} and D_{dn} depend on a weighting factor (W) (Eq. 2, 3):

$$D_{up} = \overline{WS}\sqrt{A} \quad (2)$$

$$D_{dn} = \sum_i \frac{d_i}{W_i S_i} \quad (3)$$

where, S is the average slope gradient of the upslope contributing area, A is the upslope contributing area, di and Si are the length of the flow path and the slope gradient for the i_{th} cell, respectively.

W that is intended to model the impedance to sediment fluxes was extracted in two different ways:

1) according to the linear formulation of W (Eq. 4) as a function of land use:

$$W_{lin} = 1 - n \quad (4)$$

, where n Overland Flow Manning's n Roughness Value, which depends on the land use type (Tab. 1);

2) according to the non-linear approach proposed by Gay et al. (2015) and Kalantari et al. (2017) as a function of the morphological properties and of the land use characteristics (Eq. 5):

$$W_{nl} = \frac{1}{1+e^{-0.5(x-x_0)}}\left(1 - \frac{R_i}{R_{imax}}\right) \quad (5)$$

, where RI is the roughness index dependent on the surface morphology variability (Cavalli et al., 2008; Cavalli and Marchi, 2008), RI_{max} is the highest value of RI in the study area, x_0 is the midpoint of the distribution function of RI in an area.

According to the different ways of calculation, IC distribution changes (Kalantari et al., 2017). In the considered case study, IC was calculated with both the approaches, producing two IC maps (IC_{lin} obtained implementing W_{lin}, IC_{nl} obtained implementing W_{nl}), inserted alternatively in the model for the assessment of the roads susceptible to shallow landslides.

It is worth noting that, for each trait of the road network analyzed, the value of each assigned predictor corresponded to the one of the slope immediately upstream the road trait, where a landslide, that could hit this sector, could be triggered. This is consistent with the features of the slopes where shallow landslides occurred in past in the study area. In fact, from the source area to the accumulation zone of each landslide, the failed slopes kept similar morphological and hydrological features, in

terms of slope angle, exposition, curvature and hydrological features (Bordoni et al., 2015; Persichillo et al., 2016). Maps of the predictor variables were produced for the study area at a resolution of 1 m, as the input DEM.

3.1.2 Response variable

A detailed inventory map of the road sectors affected by shallow landslides in the study area was prepared and used as response variable of the model. The inventory map of the affected road traits include all the sectors hit by the shallow landslides occurred in the study area during 27–28 April 2009, March/April 2013 and 28 February–2 March 2014rainfall events. For 2009 event, color aerial photographs at a resolution of 15 cm acquired immediately after the event were examined (Persichillo et al., 2017). For 2013 event, affected road traits were identified by visual interpretation of Pleiades satellite images with a resolution of less than 1 m (Persichillo et al., 2017). For 2014 event, slope failures and affected roads immediately after the event were detected through field surveys; the identified phenomena were mapped through a GPS tool, whose resolution is less than 2.5 m.

In the inventory map, a binary information was inserted: a value equal to 0 was assigned to the road segments not affected by a shallow landslide, while a value of 1 was assigned to each hit road trait. The resolution of the map of the response variable was set as the ones of the predictor variable (1 m).

It is worth noting that the inventory maps referred to the primary road network of the study area, composed of provincial and municipal routes. This was considered because this network contains the most affected road sectors, in terms of economic damages and indirect losses (restriction of traffic, cut-off of villages for the blockage of the road).

3.1.3 Implementation of GAM model

The data-driven methodology developed for assessing susceptible roads was based on GAM. GAM is an extension of the Generalized Linear Model (GLM), in which the linear function is replaced by an empirically fitted smooth function that allows fitting the data in the more likely functional form (Hastie and Tibshirani, 1990; Goetz et al., 2011). GAM uses a link function to relate the mean (μ) of the response variables (probability that a road sector could be hit by a landslide) and the sum of smooth functions of the predictor variables (Jia et al., 2008) (Eq. 6):

$$g(\mu) = \sum_{i=1}^{n} f_i(x_i) \quad (6)$$

, where g is the link function and the f_i are smooth function (typically splines), each dependent on a single predictor variable x_i chosen in a set of n variables $x_i \ldots x_n$. Starting from null model, each predictor variable can be included in the GAM model as linear (untransformed), non-linear (non-parametrically transformed with two equivalent degrees of freedom), or not included in the model. Predictor variables were selected through the minimization of Akaike information criterion (AIC) (Goetz et al. 2011; Persichillo et al., 2016).

GAM was implemented through 'gam' package of R software (Hastie, 2013). The adopted procedure was composed of following steps.

The first step was the application of a multicollinear analysis between the numerical predictor variables. Multicollinearity verifies when some predictor variables are linearly correlated among them to avoid redundancy that could affect the numerical stability (Farrar and Glauber 1967). The condition indexes of the matrix of the independent variables was calculated. Variables featuring such an index higher than 30 were considered not independent, thus they were excluded from the analyses to reduce collinearity (Belsley et al., 1980).

In the second step, a database formed of an equal number of road pixels affected by shallow landslides and not affected was implemented in order to avoid the over-estimation of non-landslide areas, which are much wider than landslide ones (Dai and Lee 2002; Ayalew and Yamagishi, 2005; Persichillo et al., 2016). Then, this database was subdivided into training and test sets. The training set, corresponding to 2/3 of the dataset, was used to fit the model; whereas the test set, forming of the remnant 1/3 of the dataset, was used to verify the accuracy of the model. Training and test sets were randomly selected for 100 times according to a bootstrap procedure. The most frequent predictor variables (selected 80 times at least by the bootstrap procedure) were used to build the final susceptibility model. Moreover, linear and non-linear predictors were identified according to the higher percentage of selection of each parameter.

Model forecasting capability constituted the third step of model scheme. A 100-fold repetition of holdout method for regression with a binary response, consisting of a random sub-sampling of different training and test sets, in the proportion of 2/3 for testing and 1/3 for test, was implemented. The accuracy calculated for these iterations in all training and test sets was averaged to obtain its overall value. The considered training and test sets were the ones created through the 100 bootstrap model selection. The area under the Receiver Operating Characteristic (ROC) curve (AUC) (Hosmer and Lemeshow, 2000) was computed to evaluate the model ability to discriminate affected or not road sectors, furnishing a further measure of the accuracy of the model. The AUC can take values from 0.5 (no discrimination) to 1.0 (perfect discrimination; Spitalnic, 2004). Moreover, the mean value and the bootstrap 95% confidence intervals of the 100 AUC obtained from the 100-fold bootstrap procedure for the overall accuracy of the model were calculated.

Furthermore, the 100 fitted bootstrap models were used to extend the prediction to the whole road sectors to obtain the distribution of probability. Thus, the map of the susceptibility to be hit by shallow landslides was obtained from the mean values of each bootstrap distribution of 100 probability values. Also a prediction uncertainty was associated with to each estimated probability was estimated through the calculation of by calculating the bootstrap 95% confidence intervals of the susceptibility. Different classes of probability susceptibility were created, subdividing into 4 intervals the probability values in the susceptibility map: low ($0 < p \leq 0.25$), medium–low ($0.25 < p \leq 0.50$), medium–high ($0.50 < p \leq 0.75$), high ($0.75 < p \leq 1$).

The number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) was further obtained comparing the susceptibility map with the response variable map used to build to model (Jollifee and Stephenson, 2003). For making this comparison, susceptibility values were classified as a binary variable: 1 was assigned to values higher than 0.5 (modeled pixel hit by a landslide), while 0 was assigned to values lower than 0.5 (modeled pixel not affected by a landslide).

It is worth noting that the susceptibility was calculated considering a spatial resolution of 1 m, as the input predictors, and for a buffer of 5 m from the middle of each road sector.

To assess the effect of considering or not IC in modeling the susceptibility, three models were produced and compared: Model 1) using all the predictor variables except for the IC; Model 2) considering all the predictors with IC_{lin}; Model 3) considering all the predictors with IC_{nl}.

3.2 Change in susceptibility according to different scenarios of land use

IC depends on morphological features and the land use of hillslopes, due to the presence of W factor. On the hypothesis that morphological features does not change, IC maps were created using particular land use distribution, representative of potential situations which could characterize the study area.

Besides the current scenario used for building the susceptibility models at this time, other three scenarios were considered.

The second scenario (Scenario 2) consists of the 1980 land use map, where the maximum extension of cultivated vineyards was reached (Fig. 2c). Thus, this scenario represents the possible distribution of vineyards in the case of a complete recovery of the abandoned areas since 1980s.

The third scenario (Scenario 3) corresponded to the actual scenario, with an interruption in the increase of abandoned areas without the recovery of the previously cultivated slopes. According to this, uncultivated areas completely disappear and they convert into woodlands (Fig. 5a). This scenario is consistent with the new land use management policies that were developed at the municipal level in the study area, aiming at regulating the diffusion of uncultivated areas (Rural Police Regulation, 2008; Persichillo et al., 2017).

The fourth scenario (Scenario 4) corresponded to a further increase in the abandonment of cultivated grapevines (Fig. 5b). According to this, actual uncultivated areas transform into woodlands, while further uncultivated ones develop in correspondence of actual vineyards. The slopes where abandonment was supposed are the currently cultivated ones with similar morphological features (slope angle higher than 15°) to the abandoned areas in the period 1980–2015. The increase in abandoned areas was kept equal to 22%, as that one occurred from the period 1980–2015.

Different IC scenarios were then created using these land use distributions and they were inserted in GAM model for assessing the susceptibility change of road traits in function of this parameter. Other morphological and hydrological input predictors were kept steady. The model used for these reconstructions corresponded to the one that had the best predictive performance considering the actual situation.

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-457
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 24 January 2018

Natural Hazards and Earth System Sciences
Discussions

4 Results

4.1 Map of IC reconstructed through linear and non-linear methodology

First, the distribution of IC for the actual conditions, reconstructed through the linear (IC_{lin}) and non-linear (IC_{nl}) calculation of the W factor, was analyzed (Fig. 6). This analysis was done to highlight the differences on the input IC used in modeling road susceptibility.

In the study area, IC_{lin} ranged between -7.00 and 1.75, while IC_{nl} values ranged between -4.20 and 2.23 (Fig. 7). The average value of IC distribution was -3.17 for IC_{lin} and -3.57 for IC_{nl}, while the standard deviation was similar for both the distributions (0.72 and 0.65, respectively). The map obtained with the linear implementation of W in IC calculation showed values averagely higher than the ones obtained with the non-linear W methodology in the corresponding sectors (Fig. 6).

In these analyses, IC values were classified into four classes (low, medium-low, medium-high, high), by identifying classes limits that best grouped similar values and maximized the differences between classes using the Jenks' natural breaks (Jenks, 1967), following the approach used in similar contexts by Tiranti et al. (2016), Surian et al. (2016) and Tarolli and Sofia (2016). IC_{lin} map highlighted that the northern and western parts of the catchment were characterized by medium-high and high connectivity almost totally (Fig. 6a). All the slopes with a high gradient (generally higher than 15°) presented medium-high and high connectivity features. Highest connectivity values were reached in road trenches with limited slope height (lower than 20 m) and at the bottom of hillslopes characterized by high slope angle (higher than 15-20°) and by slope height in the order of 35-70 m.

Instead, IC_{nl} map indicated lower connectivity in all the sectors of the study area (Fig. 6b). In particular, where IC_{lin} map showed a wide diffusion of slopes with medium-high and high connectivity, IC_{nl} highlighted especially medium-low and low sediment connectivity (Fig. 6b). Only few areas, close to road segments, were characterized by high connectivity (Fig. 6d). These corresponded to the road trenches characterized by a slope height lower than 20 m. It is worth noting that both reconstructions showed low and medium-low connectivity in those areas where plain areas or hillslopes with slope angle lower than 10° are present (Fig. 6).

4.2 GAM models implementation

4.2.1 Selection of the explanatory variables

Three GAM models were tested on the basis of the different set-up of the input predictors. The first phase was the selection of the variables to introduce in each model. It is important to note that all the predictors were not collinear so all these were inserted in the modelling. For each model, the variables whose selection frequency was higher than 80% in the 100-fold bootstrap procedure were selected. It was found that the variables selected were the same ones for all three models, with similar selection frequency values (Tab. 2). It is worth noting that IC was taken into account only in Model 2 and Model 3, and then consequentially selected in both these models (Tab. 2). Besides IC (having a selection frequency equal to 100% in both Models 2 and 3), the variables selected are the following (Tab. 2): SL (97%), CURV (87%), HEI (88–92%), CS (100%), TWI (85–

95%), DIST (100%), GEO (100%) in all the three models. ASP, LEN, and CA were excluded from all the models. Among these variables, only CA had a quite high frequency of selection (56-66%), but it fell under the defined threshold (Tab. 2). The selected continuous explanatory variables (all the predictors, except GEO) were distinguished into linear or non-linear, on the basis of the higher percentage of selection obtained in the bootstrap procedure. SL and HEI were chosen as linear, while CURV, CS, TWI, DIST, and IC were selected as non-linear (Tab. 2). Despite the different types of calculation of the IC implemented in Model 2 (IC_{lin}) and in Model 3 (IC_{nl}), this variable was evaluated as significant in both this model, with a frequency of selection equal to 100%. Moreover, IC was chosen as a non-linear variable in both these models IC (Tab. 2).

4.2.2 Predictive performance and susceptibility maps of the models

Model 1, that did not consider IC within the input predictors, is characterized by a fair predictive capability. In fact, AUC of the training and the test sets of this model were equal to 0.71 and 0.70, respectively (Tab. 3). AUC of the final susceptibility map produced with Model 1 is similar to those of training and test sets (0.74; Tab. 3). However, the predictive capability increased whether IC parameter was added among the predictor variables. In particular, AUC of training and test sets increased till 0.82 for Model 3, that considered also IC_{nl}. For this model, AUC of the final susceptibility map was of 0.83, with an increase of 0.09 respect to Model 1 (Tab. 3). It is important noting that a better effectiveness was reached if IC_{lin} was taken into account in a GAM model (Model 2). AUC of training and test sets reached values of 0.90, while AUC of the final susceptibility map of model 2 was of 0.94. According to Spitalnic (2004), a model with similar predictive performances can be classified as excellent.

Table 3 also highlighted very little values of standard deviation of AUC of training and test sets of each model, that maintained equal to 0.01. This confirmed the reliability of the procedure used to build up the different models.

Furthermore, the bootstrap 95% confidence intervals of AUCs were every of 0.02. This result is also confirmed by the very narrow bootstrap 95% confidence bands of ROC curves (Fig. 8a, b, c). The maps showing the bootstrap 95% confidence intervals of the probability for each road trait to be hit by a shallow landslide are illustrated in Fig. 8. As confirmed by the low values of the confidence intervals, remaining lower than 0.25, it is worth noting that the spatial variability of this probability is generally low in the entire road network of the study area for each model.

Besides, The predictive capability of the models were also evaluated by computing the values of the four indexes of a four-fold plot. TP and TN were significantly higher in Model 2 respect to Model 1 and 3, while FP and FN were significantly lower in the same model that the other two (Fig. 8d). TP and TN reached values of 90.0 and 84.8%, respectively, in Model 2, highlighting an increase of 5.1–13.7% respect to Model 3 and of 13.3–30.6% respect to Model 1. The highest effectiveness of Model 2 was confirmed also by the lowest values of FP and FN (10.0 and 15.2%, respectively), that were lower of 5.4–35.8% than the results of Model 1 and Model 3 (Fig. 8d).

The susceptibility maps for the road network extracted by GAM models are in Fig. 10.

Model 1 classified 46.9% of the road network in medium-high and high susceptibility classes. This percentage is significantly higher than those obtained for Model 2 and Model 3 (Fig. 11). The widespread of of high susceptibility areas in Model 1 explains also the high values of FP measured for this model.

Model 2 and Model 3, also considering IC within predictors, classified a lower percentage of the road network in medium-high and high susceptibility classes which are overall of 15.4% and 18.3%, respectively. (Fig. 11). The number of high susceptible road traits of Model 3 seems overestimated respect to the real situations, as demonstrated by the higher FP and FN than Model 2 (Fig. 8d). Model 2 presented a higher predictive performance than the other models, as confirmed by the quantitative indexes calculated for GAM model. It classified 15.4% of the road network of the study area in medium-high and high susceptibility classes and the remnant 85.6% in low and medium-low classes (Fig. 11). All the susceptibility maps classified as more susceptible the road sectors located below slopes of SL higher than 20°, with HEI lower than 50 m and with DIST in the range of 40–100 m. Moreover, Model 2 and Model 3 discriminated as more susceptible those road traits in correspondence of areas with medium-high and high IC, generally higher than –3, regardless of the land use which covered the slope above the road.

4.3 Susceptibility maps according to different scenarios of land use

The assessment of the predictive capabilities of the GAM models related to the actual scenarios revealed that Model 2 was the best one. This model took into account for several morphological and hydrological features of the slopes upstream the road sectors and the sediment connectivity, evaluated according to the linear modeling of IC parameter. Due to the importance of considering IC distribution in the evaluation of the routes that could be affected by shallow landslides, susceptibility scenarios were created varying IC maps input according to three defined scenarios of land use distribution hypothesized for the study area (Scenario 2, Fig. 2c; Scenario 3, Fig. 5a; Scenario 4, Fig. 5b). In fact, IC may change as a function of the change in the distribution of W factor used in the calculation of this index.

Fig. 12 illustrates the influence of different land use scenarios on IC. Its spatial distribution seemed not to be affected by land use changes presented in the considered scenarios. Thus, the connectivity of a particular hillslope kept approximately equal to the actual scenario. This was also confirmed by the mean and the standard deviation of the distribution of IC values in the study area, which remained equal to -3.20/-3.17 and 0.72/0.74, respectively.

The similar maps of IC_{lin} obtained for the different land use maps implicated that the susceptibility distribution along the road network of the study area did not change significantly for the different considered scenarios (Fig. 13). Compared to the susceptibility evaluation performed by Model 2 considering the actual scenario (Fig. 11), the differences on the percentages of the road network classified with low, medium-low, medium-high and high susceptibility by the other reconstructed scenario were negligible, ranging in the order of 0.1-0.2% (Fig. 13).

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-457
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 24 January 2018

5 Discussions

The assessment of the susceptibility of a road sector to be affected by slope instabilities is an important tool in defining the risks to the infrastructure and to the people which are engaged directly or indirectly by the same route (Bil et al., 2014). Direct and indirect economic losses induced by damages to roads provoked by landslides are widespread all over the world and are going to increase as a consequence of the growth in the triggering of landslide events caused by climate changes and economic rising income (Nemry and Demirel, 2012; Matulla et al., 2017).

For these reasons, this work aimed to develop and test a methodology able to classify, in different susceptibility classes, the traits of a road network, basing on the probability of being affected by sediments mobilized by a landslide triggered above the road. Similar approaches have been already developed and tested by different authors for particular geological/geomorphological settings (Budetta, 2004; Jaiswal et al., 2010a, 2010b, 2011; Quinn et al., 2010; Michoud et al., 2012; Bil et al., 2014; Ramesh and Anbazhagan, 2015; Pellicani et al., 2017). Instead, the proposed approach helps in filling the gaps and the limits still open in the definition of a reliable and, potentially, repeatable methodology.

First, the susceptibility model is based on a data-driven technique where several steps are carried out in order to obtain the most reliable representation of the real conditions. Data-driven models, that are based on the statistical relationships between predictors and response variables, depend strictly on the reliability of the inventories of the response variable (in this case, the sector of the road network hit by shallow landslides). These databases are used to find the statistical connections between predictor and response variables for defining the susceptibility model (Guzzetti et al., 2006; Corominas et al., 2014). Besides this limitation, data-driven are most flexible to be used at different scales of analysis (from site-specific to regional scale) and do not require a lot of data not easily to be estimated as for the physically-based models (Corominas et al., 2014). The proposed model uses GAM method applied to the susceptibility of roads to landslides for the first time. This model allows to insert some of the input predictors as non-linear variables (in this case, slope curvature, catchment slope, topographic wetness index, distance from shallow landslides source area, index of connectivity), understanding better the complex relationships which are present in an area between predisposing factors and susceptible roads (Philips, 2006; Goetz et al., 2011). Moreover, before building the model, the individuation of the most important predictor variables among the generally used predisposing factors leads to improve the knowledge about mechanisms which regulate the location of the damaged roads in such an area, avoiding for collinearity and bias that could reduce the reliability of the susceptibility estimation (Farrar and Glauber, 1967; Hosmer and Lemeshow, 1990; Bai et al., 2010). The robustness of the proposed methodology is also confirmed by the low confidence degree measured for the different susceptibility models created, testifying the reliability in implementing this scheme for the assessment of roads most susceptible to landslides (Petschko et al., 2014).

The first reconstructed susceptibility model (Model 1) takes into account for the most important predisposing factors in the study area, chosen among those morphological, hydrological and geological parameters taken into account for these analyses in different contexts by other Authors (Budetta, 2004; Jaiswal et al., 2010a, 2010b, 2011; Quinn et al., 2010; Michoud et al., 2012; Bil et al., 2014; Ramesh and Anbazhagan, 2015; Pellicani et al., 2017). The reliability of the model is quite fair, as

demonstrated by its AUC value (0.73) and by its high value of TP and TN indexes (22.3 and 45.8%, respectively). Model 1 classified as susceptible all the road segments downstream to slopes characterized by high slope gradient (> 20°), limited height (< 50 m) and with shallow landslides triggering zones located very close to the road network (40-100 m). These settings are very widespread in the entire study area (Bordoni et al., 2015; Persichillo et al., 2016, 2018), but these particular features are not enough to discriminate more accurately those routes where damages provoked by sediments mobilized by shallow landslides are probable.

Tarolli and Sofia (2016) and Persichillo et al. (2018) analyzed two different hilly and mountainous catchments, located in western USA and north-eastern Oltrepò Pavese, respectively, in order to highlight potential connections between a road network and the sediment delivery. Both the works quantified the sediment connectivity using the index of connectivity (IC), that allows evaluating the potential connection between hillslopes and features which act as targets for transported sediments based only on of the morphological and topographical characteristics and the vegetation cover of a territory. These works highlighted that the segments of the road network, which can act as a storage area for the sediments mobilized by a phenomenon upstream to the road, are those ones located in correspondence of zones characterized by high IC values. This aspect testifies how slope instability phenomena can actively deliver sediment to particular portions of a road network, producing also damages provoked by the impact of the mobilized materials with the infrastructure (Sidle et al., 2014; Klose et al., 2015).

Starting from this observation, also IC was inserted in the model for the evaluation of the susceptibility of the roads to be affected by shallow landslides. Other two models were created, differing each other for the type of IC used. Model 2 uses IC_{lin} calculated according to the method proposed by Cavalli et al. (2013), where the W factor in the model is evaluated in a linear way. Model 3 uses IC_{nl}, calculated by means also of a W factor evaluated in a non-linear way and in relation also to the both surface roughness and land use properties of a territory (Fryirs et al., 2007; Cavalli et al., 2008; Cavalli and Marchi, 2008; Gay et al., 2016; Kalantari et al., 2017). In these terms, both IC_{lin} and IC_{nl} represent a structural connectivity depending on the morphological and land use attributes of a territory (Borselli et al., 2008; Cavalli et al., 2013; Crema and Cavalli, 2015).

Models that consider also sediment connectivity have a higher predictive performance than model 1. This is testified by a high AUC (0.94 for Model 2 and 0.83 for Model 3) and by higher values of TP and TN (till 90.0% and 84.8%, respectively). Moreover, FP and FN of both these models are lower than Model 1 (till 10.0 and 15.2 respectively).

The susceptibility maps produced through Model 2 and Model 3 identify the road sectors characterized by the highest values of IC (IC higher than –3) as the most susceptible. These conditions are measured in several routes regardless of the land cover present in the slope upstream the road. Among these models, Model 2, that consider IC calculated through the linear way, performs better than Model 3. Non-linearly reconstructed IC identified less areas with high connectivity than IC_{lin}. Thus, the estimated probability to be affected by sediment impacts is reduced in these road traits. IC_{nl} is more representative of the sediment connectivity in lowland environments, where the connectivity is driven also by other factors (such as the amount of surface water runoff) together with the morphological features of the hillslopes (Fryirs et al., 2007; Gay et al., 2016; Kalantari et al., 2017). Its application in geomorphological settings characterized by a predominant hilly or mountainous morphology, such as the considered catchment, can implicate an underestimation on the connectivity or disconnectivity of the sediments,

influencing also the correct assessment of the sectors of an infrastructure threatened by the material mobilized after a triggering event.

The comparison of the best susceptibility model (Model 2) with the distribution of real case of road sectors damaged by sediments mobilized by shallow landslides triggered upstream the road has confirmed an excellent predictive performance of this model. This model allows to identify correctly both the road sectors hit by the accumulation zones of roto-translational shallow landslides triggered in the trenches present in halfway roads (B2 type) and the road traits affected by the materials mobilized by shallow landslides triggered in the slopes upstream the routes in correspondence of cultivated or abandoned hillslopes (Fig. 14). This reveals the suppleness of the methodology to estimate in a reliable way the most susceptible sectors of a road network also in the case of sediment source areas, represented by slope instabilities with different features.

Nevertheless, Model 2 classifies wrongly some road sectors in the study area, as testified by the 15.2% of FN and of 10% of FP cases. FN are mostly located in few pixels, sometimes close to other road traits identified with medium-high or very-high susceptibility. These situations could be linked to local peculiar factors, which may affect road susceptibility, that are not completely described by the input predictors chosen for the model. On the contrary, FP cases correspond to road segments where high susceptibility (higher than 0.5) was estimated. These sectors are mostly located near to traits already affected by shallow landslides materials in past events (Fig. 15). They are in a buffer of less than 250 m, in particular between 50 and 200 m, respect to sectors hit in past, and they present morphological and connectivity features similar to threatened traits. Hence, they could represent sectors which could be affected by future events occurred in the same study area, whether the settings of this zone and the triggering conditions will keep similar to past events. In these terms, susceptibility map obtained from Model 2 is useful in determining accurately the susceptible sectors of a road network, furnishing an important tool for the management of the hazard and for sketching policies of risk reduction out.

Due to the importance of sediment connectivity on the model capability, scenarios of susceptibility were reconstructed, through Model 2, starting from different IC maps were obtained considering particular land use distributions. In fact, changes in land use cause are represented by changes in W parameter of the IC calculation, provoking a potential variation in the connectivity distribution. The distribution of susceptibility and of the roads most probably affected by shallow landslides do not change significantly from the actual situation for the three different modeled scenarios (recovery of all cultivated vineyards, break on the abandonment, further increase of the abandoned areas). This is due to the similar values of W (0.6-0.8) characterizing the most widespread land covers of the study area, which thus induce to a limited change in IC value passing from a land use class to another one. Instead, it is worth noting that changes in land use distribution could have effects also on the physical morphology of the hillslopes (Fu et al., 2006; Tarolli et al., 2015). For example, the recovery of the cultivation of grapevines in a slope could lead to the development of a drainage system of the superficial and of the shallow waters and to modification on the slope morphology for the implantation of the vineyards. While the abandonment of previously cultivated vineyards induces changes in flow direction and regulation, with direct consequences on sediment production and delivery (Cevasco et al., 2014; Lieskovsky and Kenderessy, 2014; Tarolli et al., 2014; Prosdocimi et al., 2016). These actions could influence the movement of the rainwater and of the sediment mobilized by runoff, then reduce the connectivity and also the potential

susceptibility of a road located downstream. Hence, more detailed scenarios of susceptibility changes in relation to land use changes will take into account also for the morphological modifications linked to these changes, using also a higher resolution DEM (less than 1 m).

6 Conclusions

In this work, a non-linear data-driven approach, based on GAM, was developed for the evaluation of the susceptible road sectors of a network that could be affected by the sediments delivered from shallow landslides occurred upstream. The methodology assessed also the role of the sediment connectivity on susceptibility estimation, by the implementation of the index connectivity calculated according to a linear or a non-linear approach.

Besides the use of an inventory of road damages referred only to three triggering events occurred from 2009 to 2014, the random partition of the entire dataset in two parts (training and test subsets), within a 100-fold bootstrap procedure, allowed to select the most significant predisposing variables. This provided a better description of the occurrence and distribution of the road sectors potentially susceptible to damages induced by shallow landslides.

The best predictive capability was reached by a model which took into account also the index of connectivity, calculated according to a linear way. This index well represented the rates of connectivity and disconnectivity in the studied catchment, in relation to its morphology (steep slopes, narrow valleys) and the land uses (vineyards, abandoned areas, woodlands). Most susceptible road traits resulted in the ones located below steep slopes with a limited height (lower than 50 m), where sediment connectivity is high, regardless of the land use which covered the slope above the road.

Different scenarios of land use were implemented in order to estimate possible changes in road susceptibility. Land use classes of the study area were characterized by similar effects on connectivity features, thus the index of connectivity did not change significantly with a consequent leakage of variations also on the susceptibility of the road networks. Larger effects on sediment connectivity could be induced by modifications in the morphology of the slopes (e.g. drainage system, modification of the slope angle) provoked by the abandonment or by the recovery of cultivations. Then, this could have effects on the sediment delivery and also on the susceptibility of a road to be hit by sediments mobilized upstream.

The presented methodology allows identifying the most susceptible road sectors that could be hit by sediments delivered by landslides in a robust and reliable way. This tool can represent a fundamental starting point for improving the land management of the slopes where the source areas of the sediments could develop, in order to reduce the damages to the infrastructure and the related risks and economic losses. Moreover, the results of the susceptibility analysis can give asset managers indispensable information on the relative criticality of the different road sectors, thereby allowing attention and economic budgets to be shifted towards the most critical assets, where structural and non-structural mitigation measures could be implemented.

Furthermore, thanks to the flexibility of the model in the selection of the predictors, the proposed model can be applied to areas with different geological, geomorphological and land use features, identifying the most important predisposing factors peculiar of each catchment. This method can be also implemented to areas characterized by much larger catchments than the ones

analyzed herein, with the only limit of the availability of high-resolution DEMs and of computational resources. Moreover, the methodology can be applied for estimating the susceptibility and the risks related to landslides affecting other fundamental facilities, such as railways, gas/oil pipelines, power lines.

Authors contribution

Massimiliano Bordoni analyzed the data, developed the methodological approach and prepared the manuscript; Maria Giuseppina Persichillo helped in the development of the methodology and in the interpretation of the results; Claudia Meisina helped in the interpretation of the results and provided guidance and support throughout the research process; Stefano Crema and Marco Cavalli helped in building the connectivity scenarios and in supporting their interpretation; Carlotta Bartelletti, Yuri Galanti and Michele Barsanti helped in the development of the data-driven approach and in the interpretation of the results; Roberto Giannecchini and Giacomo D'Amato Avanzi collaborated in writing the manuscript and in the comprehension of the analyses.

References

Ayalew, L. and Yamagishi, H.: The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan, Geomorphology 65, 12–31, doi:10.1016/j.geomorph.2004.06.010, 2005.

Bai, S. B., Wang, J., Lu, G. N., Zhou, P. G., Hou, S. S. and Xu, S. N.: Gis-based logistic regression fro landslide susceptibility mapping of the Zhongxian segment in the Three Gorges area, China, Geomorphology 115, 1-2, 23-31, doi: 10.1016/j.geomorph.2009.09.025, 2010.

Bathurst, J. C., Burton, A. and Ward, T.J.: Debris flow run-out and landslide sediment delivery model test, J. Hydraul. Eng., 123, 5, 410-419, doi:10.1061/(ASCE)0733-9429(1997)123:5(410), 1997.

Begueria, 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.

Belsley, D. A., Kuh, E. and Welsch, R. E. (Eds.): Regression diagnostics: identifying influential data and sources of collinearity, John Wiley and Sons, New York, USA, 1980.

Bil, M., Kubecek, J. and Andrasik, R.: An epidemiological approach to determining the risk of road damage due to landslides, Nat. Hazards 73, 1323–1335, doi:10.1007/s11069-014-1141-4, 2014.

Bordoni, M., Meisina, C., Valentino, R., Bittelli, M. and Chersich, S.: Site-specific to local scale shallow landslides triggering zones assessment using TRIGRS, Nat. Hazards Earth Syst. Sci. 15, 1025–1050, doi:10.5194/nhess-15-1025-2015, 2015.

Borselli, L., Cassi, P. and Torri, D.: Prolegomena to sediment and flow connectivity in the landscape: a GIS and field numerical assessment, Catena 75, 268–277, doi:10.1016/j.catena.2008.07.006, 2008.

Brenning, A., Schwinn, M., Ruiz-Paez, A. P. and Muenchow, J.: Landslide susceptibility near highways is increased by 1 order of magnitude in the Andes of southern Ecuador, Loja province, Nat. Hazards Earth Syst. Sci., 15, 45–57, doi:10.5194/nhess-15-45-2015, 2015.

Budetta, P.: Assessment of rockfall risk along roads, Nat. Hazards Earth Syst. Sci. 4, 1, 71–81, doi:10.5194/nhess-4-71-2004, 2004.

Catani, F., Lagomarsino, D, Segoni, S. and Tofani, V.: Landslide susceptibility estimation by random forests technique: sensitivity and scaling issues, Nat. Hazards Earth Syst. Sci. 13, 2815–2831, doi:10.5194/nhess-13-2815-2013, 2013.

Cavalli, M. and Marchi, L.: Characterisation of the surface morphology of an alpine alluvial fan using airborne LiDAR, Nat. Hazards Earth Syst. Sci. 8, 323–333, doi:10.5194/nhess-8-323-2008, 2008.

Cavalli, M., Tarolli, P., Marchi, L. and Dalla Fontana, G.: The effectiveness of airborne LiDAR data in the recognition of channel-bed morphology, Catena 73 249–260, doi:10.1016/j.catena.2007.11.001, 2008.

Cavalli, M., Trevisani, S., Comiti, F. and Marchi, L.: Geomorphometric assessment of spatial sediment connectivity in small alpine catchments, Geomorphology 188, 31–41, doi:10.1016/j.geomorph.2012.05.007, 2013.

Cevasco, A., Pepe, G. and Brandolini, P.: The influences of geological and land-use settings on shallow landslides triggered by an intense rainfall event in a coastal terraced environment, Bull. Eng. Geol. Environ. 73, 859–875, doi:10.1007/s10064-013-0544-x, 2014.

Chau, K. T., Sze, Y. L., Fung, M. K., Wong, W. Y., Fong, E. L. and Chan, L. C. P.: Landslide hazard analysis for Hong Kong using landslide inventory and GIS, Comp. Geosci. 30, 429-443, doi:10.1016/j.cageo.2003.08.013, 2004.

Chen, Z. and Wang, J.: Landslide hazard mapping using logistic regression model in Mackenzie Valley, Canada, Nat. Hazards 42, 75–89, doi:10.1007/s11069-006-9061-6, 2007.

Corominas, J., Van Westen, C., Frattini, P., Cascini, L., Malet, J. P., Fotopoulou, S., Catani, F., Van Den Eeckhaut, M., Mavrouli, O., Agliardi, F., Pitilakis, K., Winter, M. G., Pastor, M., Ferlisi, S., Tofani, V., Hervas, J., and Smith, J. T.: Recommendations for the quantitative analysis of landslide risk, Bull. Eng. Geol. Environ., 73, 209–263, doi:10.1007/s10064-013-0538-8 2014.

Crema S., Cavalli M.: SedInConnect: A stand-alone, free and open source tool for the assessment of sediment connectivity, Comp. Geosci., 111, 39-45, doi:10.1016/j.cageo.2017.10.009, 2018.

Cruden, D. M. and Varnes, D. J.: Landslide types and processes, in: Landslides: investigation and mitigation, edited by: Turner, A. K. and Schuster, R. L., National Academy Press, Washington, D.C., 36–75, 1996.

Dai, F. C., Lee, C. F. and Ngai, Y. Y.: Landslide risk assessment and management: an overview, Eng. Geol. 64, 65–87, doi:10.1016/S0013-7952(01)00093-X, 2002.

Dai, F. C. and Lee, C. F.: Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong, Geomorphology 42, 213–228, doi:10.1016/S0169-555X(01)00087-3, 2002.

D'Amato Avanzi, G., Galanti, Y., Giannecchini, R. and Puccinelli, A.: Fragility of territory and infrastructures resulting from rainstorms in Northern Tuscany (Italy), in: Landslide science and practice vol. 6., edited by: Margottini, C., Canuti, P. and Sassa, K., Springer-Verlag Berlin Heidelberg, 239-246, 2013.

Fan, L., Lehmann, P., McArdell, B. and Or, D.: Linking rainfall-induced landslides with debris flows runout patterns towards catchment scale hazard assessment, Geomorphology 280, 1–15, doi:10.1016/j.geomorph.2016.10.007, 2017.

Fannin, R. and Wise, M.: An empirical-statistical model for debris flow travel distance, Can. Geotech. J. 38, 5, 982–994, doi:10.1139/t01-030, 2001.

Farrar, D. E. and Glauber, R. R.: Multicollinearity in regression analysis: the problem revisited, Rev. Econ. Stat. 49, 92–107, doi:10.2307/1937887, 1967.

Fasolini, D.: La cartografia dell'uso e copertura del suolo: uno strumento per rilevare il cambiamento del territorio lombardo, Regione Lombardia e ERSAF, 76–87, 2014.

Fathani, T. F., Legono, D. and Karnawati, D.: A numerical model for the analysis of rapid landslide motion, Geotech. Geol. Eng., 35, 2253–2268, doi:10.1007/s10706-017-0241-9, 2017.

Foerster, S., Wilczok, C., Brosinsky, A. and Segl, K.: Assessment of sediment connectivity from vegetation cover and topography using remotely sensed data in a dryland catchment in the Spanish Pyrenees, J. Soils Sediments 14, 1982–2000, doi:10.1007/s11368-014-0992-3, 2014.

Fryirs, K. A., Brierley, G. J., Preston, N. J. and Kasai, M.: Buffers, barriers and blankets: the (dis)connectivity of catchment-scale sediment cascades, Catena 70, 49–67, doi:10.1016/j.catena.2006.07.007, 2007.

Fu, B. J., Zhang, Q. J., Chen, L. D., Zhao, W. W., Gulinck, H., Liu, G. B., Yang, Q. K. and Zhu, Y. G.: Temporal change in land use and its relationship to slope degree and soil type in a small catchment on the Loess Plateau of China, Catena, 65, 41-48, doi:10.1016/j.catena.2005.07.005, 2006.

Galve, J. P., Cevasco, A., Brandolini, P. and Soldati, M.: Assessment of shallow landslide risk mitigation measures based on land use planning through probabilistic modelling, Landslides 12, 101–114, doi:10.1007/s10346-014-0478-9, 2015.

Gariano, S. L. and Guzzetti, F.: Landslides in a changing climate, Earth Sci. Rev. 162, 227-252, doi:10.1016/j.earscirev.2016.08.011, 2016.

Gay, A., Cerdan, O., Mardhel, V. and Desmet, M.: Application of an index of sediment connectivity in a lowland area, J. Soils Sediments 16, 280–293, doi:10.1007/s11368-015-1235-y, 2015.

Glade, T.: Landslide occurrence as a response to land use change: a review of evidence from New Zealand, Catena 51, 297–314, doi:10.1016/S0341-8162(02)00170-4, 2003.

Goetz, J. N, Guthrie, R. H. and Brenning, A.: Integrating physical and empirical landslide susceptibility models using generalized additive models, Geomorphology 129, 376–386, doi:10.1016/j.geomorph.2011.03.001, 2011.

Guzzetti, F., Galli, M., Reichenbach, P., Ardizzone, F. and Cardinali, M.: Landslide hazard assessment in the Collazzone area, Umbria, Central Italy, Nat. Hazards Earth Syst. Sci. 6, 115–131, doi:10.5194/nhess-6-115-2006, 2006.

Hastie, T. J. and Tibshirani, R. (Eds.): Generalized additive models, Chapman & Hall, London, England, 1990.

Hastie, T.: gam: Generalized Additive Models, R package version 1.08, http://CRAN.R-project.org/package=gam, last access: 29 August 2013.

Hosmer, D. W. and Lemeshow, S.: Applied logistic regression, Wiley, New York, USA, 2000.

Hungr, O.: A model for the runout analysis of rapid flow slides, debris flows, and avalanches, Can. Geotech. J. 32, 610–623, doi:10.1139/t95-063, 1995.

Jaiswal, P., Van Westen, C. J. and Jetten, V.: Quantitative landslide hazard assessment along a transportation corridor in southern India, Eng. Geol. 116, 3, 236–250, doi:10.1016/j.enggeo.2010.09.005, 2010a.

Jaiswal, P., Van Westen, C. J. and Jetten, V.: Quantitative assessment of direct and indirect landslide risk along transportation lines in southern India, Nat. Hazards Earth Syst. Sci. 10, 6, 1253–1267, doi:10.5194/nhess-10-1253-2010, 2010b.

Jaiswal, P., Van Westen, C. J. and Jetten, V.: Quantitative assessment of landslide hazard along transportation lines using historical records, Landslides 8, 3, 279–291, doi:10.1007/s10346-011-0252-1, 2011.

Jenks, G. F.: The data model concept in statistical mapping, Int. Year. Cart. 7, 186–190, 1967.

Jia, G., Yuan, T., Liu, Y. and Zhang, Y.: A static and dynamic factors-coupled forecasting model or regional rainfall induced landslides: a case study of Shenzhen, Sci. China Ser. E. 51 164–175, 2008.

Jollifee, I. T. and Stephenson, D. B. (Eds.): Forecast verification: A practitioner's guide in atmospheric science, John Wiley and Sons, New York, USA; 2003.

Kalantari, Z., Cavalli, M., Cantone, C., Crema, S. and Destouni, G.: Flood probability quantification for road infrastructure: Data-driven spatial-statistical approach and case study applications, Sci. Tot. Environ. 581-582, 386-398, doi:10.1016/j.scitotenv.2016.12.147, 2017.

Klose, M., Damm, B. and Terhorstet, B.: Landslide cost modeling for transportation infrastructures: a methodological approach, Landslides 12, 321–334, doi:10.1007/s10346-014-0481-1, 2015.

Kritikos, T. and Davies, T.: Assessment of rainfall-generated shallow landslide/debris-flow susceptibility and runout using a GIS-based approach: application to western Southern Alps of New Zealand, Landslides 12, 1051-1075, doi:10.1007/s10346-014-0533-6, 2015.

Lieskovsky, L. and Kenderessy, P.: Modelling the effect of vegetation cover and different tillage practices on soil erosion in vineyards: a case study in Vrable (Slovakia) using WATEM/SEDEM, Land Degrad. Dev. 25, 188–196, doi:10.1002/ldr.2162 2014.

Lopez-Vicente, M., Poesen, J., Navas, A. and Gaspar, L.: Predicting runoff and sediment connectivity and soil erosion by water for different land use scenarios in the Spanish Pre-Pyrenees, Catena 102, 62-73, doi:10.1016/j.catena.2011.01.001, 2013.

Lopez-Vicente, M., Nadal-Romero, E. and Cammeraat, E. L. H.: Hydrological connectivity does change over 70 years of abandonment and afforestation in the Spanish Pyrenees, Land Degrad. Dev., doi:10.1002/ldr.2531, 2016.

Martinovic, K., Gavin, K. and Reale, C.: Development of a landslide susceptibility assessment for a rail network, Eng. Geol. 215, 1-9, doi:10.1016/j.enggeo.2016.10.011, 2016.

Matulla, C., Hollsi, B., Andre1, K., Gringinger, J., Chimani, B., Namyslo, J., Fuchs, T., Auerbach, M., Herrmann, C., Sladek, B., Berghold, H., Gschier, R. and Eichinger-Vill, E.: Climate Change driven evolution of hazards to Europe's transport infrastructure throughout the twenty-first century, Theor. Appl. Climatol., doi:10.1007/s00704-017-2127-4, 2017.

Michoud, C., Derron, M. H., Horton, P., Jaboyedoff, M., Baillifard, F. J., Loye, A., Nicolet, P., Pedrazzini, A. and Queyrel, A.: Rockfall hazard and risk assessments along roads at a regional scale: example in Swiss Alps, Nat. Hazards Earth Syst. Sci. 12, 3, 615–629, doi:10.5194/nhess-12-615-2012 ,2012.

Petschko, H., Brenning, A., Bell, R., Goetz, J. and Glade, T.: Assessing the quality of landslide susceptibility maps – case study Lower Austria, Nat. Hazards Earth Syst. Sci. 14, 95–118, doi:10.5194/nhess-14-95-2014 2014.

Muenchow, J., Brenning, A., and Richter, M.: Geomorphic process rates of landslides along a humidity gradient in the tropical Andes, Geomorphology, 139, 271–284, doi:10.1016/j.geomorph.2011.10.029 2012.

Nemry, F. and Demirel, H.: Impacts of Climate Change on transport: a focus on road and rail transport infrastructures. Publications Office of the European Union, Luxembourg, Luxembourg, 93 pp., 2012.

Pastor, M., Blanc, T., Haddad, B., Petrone, S., Sanchez Morles, M., Drempetic, V., Issler, D., Crosta, G. B., Cascini, L., Sorbino, G. and Cuomo, S.: Application of a SPH depth-integrated model to landslide run-out analysis, Landslides 11, 793–812, doi:10.1007/s10346-014-0484-y, 2014.

Pellicani, R., Argentiero, I. and Spilotro, G.: GIS-based predictive models for regional-scale landslide susceptibility assessment and risk mapping along road corridors, Geomat. Nat. Haz. Risk, doi:10.1080/19475705.2017.1292411, 2017.

Persichillo, M. G., Bordoni, M., Meisina, C., Bartelletti, C., Barsanti, M., Giannecchini, R., D'Amato Avanzi, G., Galanti, Y., Cevasco, A., Brandolini, P. and Galve, J. P.: Shallow landslides susceptibility assessment in different environments, Geomat. Nat. Haz. Risk, doi:10.1080/19475705.2016.1265011, 2016.

Persichillo, M. G., Bordoni, M. and Meisina, C.: The role of land use changes in the distribution of shallow landslides, Sci. Total Environ. 574, 924–937, doi:10.1016/j.scitotenv.2016.09.125, 2017.

Persichillo, M. G., Bordoni, M., Cavalli, M., Crema, S. and Meisina, C.: The role of human activities on sediment connectivity of shallow landslides, Catena 160, 261-274, doi:10.1016/j.catena.2017.09.025, 2018.

Phillips, J. D.: Sources of nonlinearity and complexity in geomorphic systems, Prog. Phys. Geog 27, 1–23, doi:10.1191/0309133303pp340ra, 2003.

Phillips, J. D.: Evolutionary geomorphology: thresholds and nonlinearity in landform response to environmental change, Hydrol. Earth Syst. Sci. 10, 731–742, doi:10.5194/hess-10-731-2006, 2006.

Prosdocimi, M., Cerdà, A. and Tarolli, P.: Soil water erosion on Mediterranean vineyards: A review, Catena 141, 1-21, doi:10.1016/j.catena.2016.02.010, 2016.

Quinn, P., Beven, K., Chevallier, P. and Planchon, O.: The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models, Hydrol. Processes. 5, 59–79, doi:10.1002/hyp.3360050106, 1991.

Quinn, P. E., Hutchinson, D. J., Diederichs, M. S. and Rowe, R. K.: Regional-scale landslide susceptibility mapping using the weights of evidence method: an example applied to linear infrastructure, Can. Geotech. J. 47, 905–927, doi:10.1139/T09-144, 2010.

Ramesh, V. and Anbazhagan, S.: Landslide susceptibility mapping along Kolli hills Ghat road section (India) using frequency ratio, relative effect and fuzzy logic models, Environ. Earth Sci. 73, 8009–8021, doi:10.1007/s12665-014-3954-6 2015.

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. Manag. 54, 1372–1384, doi:10.1007/s00267-014-0357-0 2014.

Rural Police Regulation: Regolamento di Polizia Rurale, Comune di Canneto Pavese, Italy, 17 pp., 2008.

Salvati, P., Bianchi, C., Fiorucci, F., Giostrella, P., Marchesini, I. and Guzzetti, F.: Perception of flood and landslide risk in Italy: a preliminary analysis, Nat. Hazards Earth Syst. Sci. 14, 2589–2603, doi:10.5194/nhess-14-2589-2014, 2014.

Seibert, J,. Stendahl, J. and Sorensen, R.: Topographical influences on soil properties in boreal forests, Geoderma, 141, 139–148, doi:10.1016/j.geoderma.2007.05.013, 2007.

Sidle, R. C. and Ochiai, H.: Landslides: Processes, prediction, and land use,Water Resources Monograph, AGU, Washington D.C., 2006.

Sidle, R. C., Ghestem, M. and Stokes, A.: Epic landslide erosion from mountain roads in Yunnan, China – challenges for sustainable development, Nat. Hazards Earth Syst. Sci. 14, 3093–3104, doi:10.5194/nhess-14-3093-2014, 2014.

Spitalnic, S.: Test properties 2: likelihood ratios, Bayes' formula, and receiver operating characteristic curves, Hosp. Physician 40, 53–58, 2004.

Surian, N., Righini, M., Lucia, A., Nardi, L., Amponsah, W., Benvenuti, M., Borga, M., Cavalli, M., Comiti, F., Marchi, L., Rinaldi, M. and Viero, A.: Channel response to extreme floods: Insights on controlling factors from six mountain rivers in northern Apennines, Italy, Geomorphology 272, 78–91, doi:10.1016/j.geomorph.2016.02.002, 2016.

Tarolli, P., Preti, F. and Romano, N.: Terraced landscapes: from an old best practice to a potential hazard for soil degradation due to land abandonment, Anthropocene 6, 10–25, doi:10.1016/j.ancene.2014.03.002, 2014.

Tarolli, P., Sofia, G., Calligaro, S., Prosdocimi, M., Preti, F. and Dalla Fontana, G.: Vineyards in terraced landscapes: New opportunities from LIDAR data, Land Degrad. Develop. 26, 92–102, doi:10.1002/ldr.2311, 2015.

Tarolli, P. and Sofia, G.: Human topographic signatures and derived geomorphic processes across landscapes, Geomorphology 255, 140–161, doi:10.1016/j.geomorph.2015.12.007, 2016.

Tiranti, D., Cavalli, M., Crema, S., Zerbato, M., Graziadei, M., Barbero, S., Cremonini, R., Silvestro, C., Bodrato, G. and Tresso, F.: Semi-quantitative method for the assessment of debris supply from slopes to river in ungauged catchments, Sci. Total Environ. 554–555, 337–348, doi:10.1016/j.scitotenv.2016.02.150, 2016.

Van Westen, C. J., Asch T. W. J. and Soeters, R.: Landslide hazard and risk zonation—why is it still so difficult?, Bull. Eng. Geol. Environ. 65, 67–184, doi:10.1007/s10064-005-0023-0, 2006.

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

Varnes, D. J. (Ed.): Landslide hazard zonation – a review of principles and practice, UNESCO, Paris, France, 1984.

Vercesi, P. and Scagni, G.: Osservazioni sui depositi conglomeratici dello sperone collinare di Stradella, Rendiconti della Società Geologica Italiana, 7, 23–26, 1984.

Zezere, J. L., Oliveira, S. C. and Garcia, R. A. C. and Reis, E.: Landslide risk analysis in the area North of Lisbon (Portugal): evaluation of direct and indirect costs resulting from a motorway disruption by slope movements, Landslides 4, 123–136, doi:10.1007/s10346-006-0070-z, 2007.

Zevenbergen, L. W. and Thorne, C. R.: Quantitative analysis of land surface topography. Earth Surf. Processes Landforms 12, 47–56, doi:10.1002/esp.3290120107, 1987.

Zizioli, D., Meisina, C., Valentino, R., and Montrasio, L.: Comparison between different approaches to modeling shallow landslide susceptibility: a case history in Oltrepo Pavese, Northern Italy, Nat. Hazards Earth Syst. Sci., 13, 559–573, doi:10.5194/nhess-13-559-2013, 2013.

Table 1: Overland flow Manning's n Roughness Values assigned to each class of land use maps available for the calculation of W_{lin} factor.

Land use classes	Manning's n (-)
Woodlands	0.40
Uncultivated areas	0.35
Grasslands	0.25
Orchards/Arable areas/Vineyards	0.20
Bare soil	0.05
Urban areas	0.02

Table 2: Frequencies (in %) of explanatory variables (both linear and non-linear) selected by 100-fold bootstrap procedure. The explanatory variables selected by 100-fold bootstrap procedure with an absolute frequency greater or equal than 80% are highlighted in bold red. In the brackets, the frequencies (in %) of selection of each variable as linear or non-linear is shown. The underlined number corresponded to the frequency of the selected function connected to each variable. SL: slope angle; ASP: slope aspect; CURV: slope curvature; LEN: slope length; HEI: slope height; CA: Catchment area; CS: Catchment slope; TWI: topographic wetness index; DIST: distance from the source area of a shallow landslide; GEO: bedrock geology; IC: index of connectivity.

Model	SL	ASP	CURV	LEN	HEI	CA	CS	TWI	DIST	GEO	IC
1 (Lin.-Not Lin.)	**97 (95-2)**	2 (0-2)	**87 (41-46)**	36 (36-0)	**88 (45-43)**	56 (56-0)	**100 (0-100)**	**85 (14-71)**	**100 (0-100)**	**100**	-
2 (Lin.-Not Lin.)	**97 (95-2)**	18 (12-6)	**87 (41-46)**	19 (19-0)	**88 (45-43)**	66 (66-0)	**100 (0-100)**	**85 (16-69)**	**100 (0-100)**	**100**	**100 (5-95)**

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-457
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 24 January 2018

Natural Hazards and Earth System Sciences
Discussions

3 (Lin.-Not Lin.)	**97 (95-2)**	18 (11-7)	**87 (41-46)**	19 (19-0)	**92 (53-39)**	65 (65-0)	**100 (0-100)**	**95 (12-83)**	**100 (0-100)**	**100**	**100 (2-98)**

Table 3: Mean and standard deviation of accuracy for the training sets, the test sets and the final application of the model to the entire study area.

Model	Mean accuracy of training sets (-)	Standard deviation of accuracy on training sets (-)	Mean accuracy on test sets (-)	Standard deviation of accuracy on test sets (-)	Mean AUC of the model (-)	95 % confidence interval of AUC of the model (-)
1	0.71	0.01	0.70	0.01	0.74	0.73-0.75
2	0.90	0.01	0.90	0.01	0.94	0.93-0.95
3	0.82	0.01	0.82	0.01	0.83	0.82-0.84

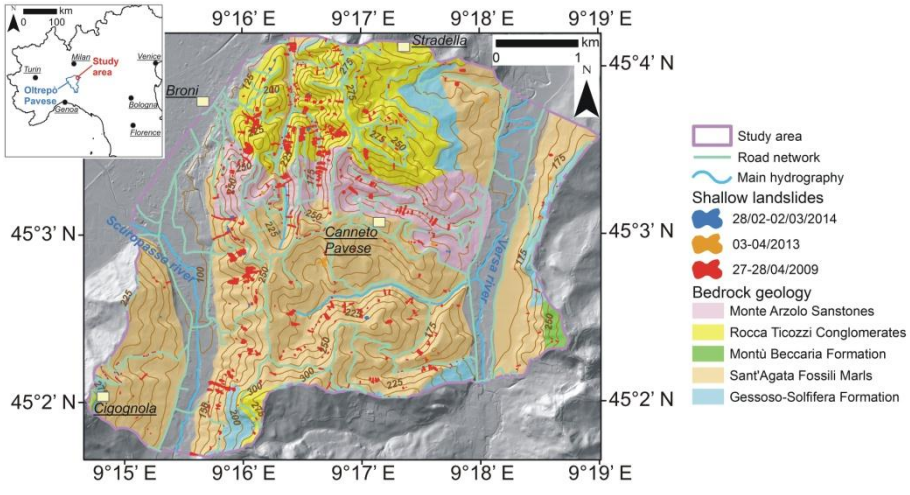


Figure 1: Geological setting and shallow landslides distribution of the study area.

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-457
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 24 January 2018

Natural Hazards and Earth System Sciences
Discussions

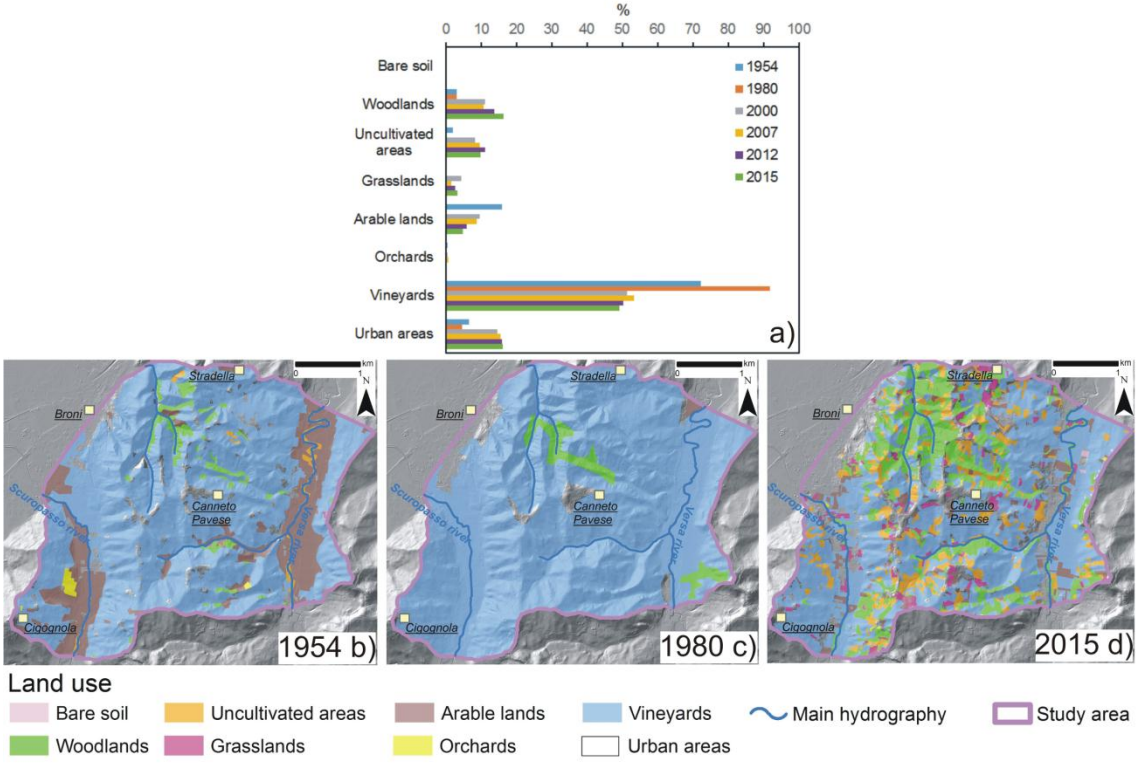


Figure 2: Land use distribution and land use changes in the period 1954-2015: a) percentage of the area occupied by each land use class during the analyzed period; b) land use distribution in 1954; c) land use distribution in 1980; d) land use distribution in 2015. Land use maps were provided by the Lombardy Region and shared as part of the Infrastructure for Spatial Information in Lombardy (IIT) via the Geoportal (Lombardy Region Geoportal: http://www.cartografia.regione.lombardia.it/geoportale, last access: 11 December 2017). The detailed information regarding the method to realize these maps are available in Fasolini (2014).

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-457
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 24 January 2018

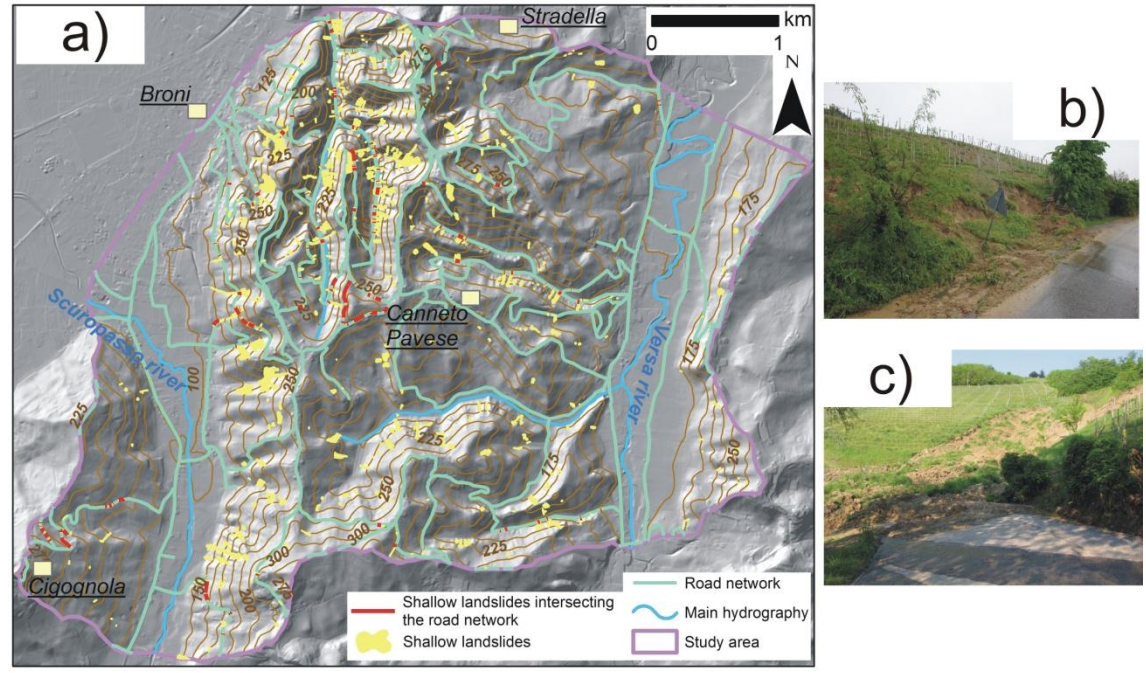


Figure 3: a) Primary road network of the study area, with the shallow landslides events occurred between 2009 and 2014 which affected these routes. This road network is composed of provincial and municipal roads (available from: Administration of Pavia Province and Infrastructure for Spatial Information in Lombardy (Lombardy Region Geoportal: http://www.cartografia.regione.lombardia.it/geoportale, last access: 11 December 2017). b) A shallow landslide (B2 type), triggered in correspondence of the road trench upstream the route, that blocked the route. c) A shallow landslide triggered in a slope cultivated with vineyards, whose mobilized materials destroyed completely a road trait downstream.

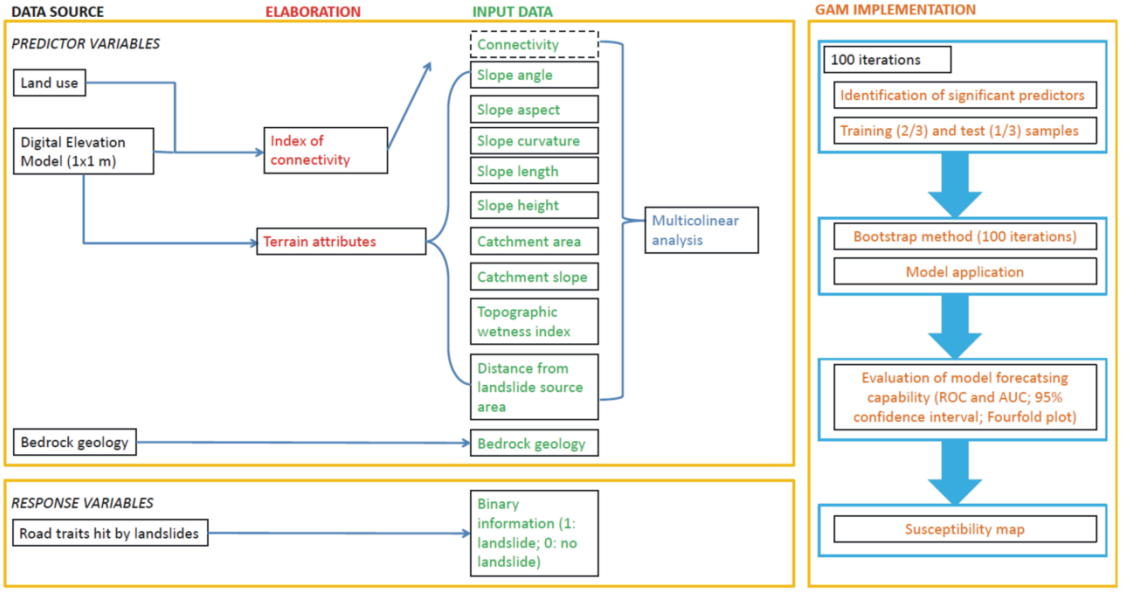


Figure 4: Flow-chart containing the scheme for the implementation of the proposed model.

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-457
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 24 January 2018

Natural Hazards and Earth System Sciences
Discussions

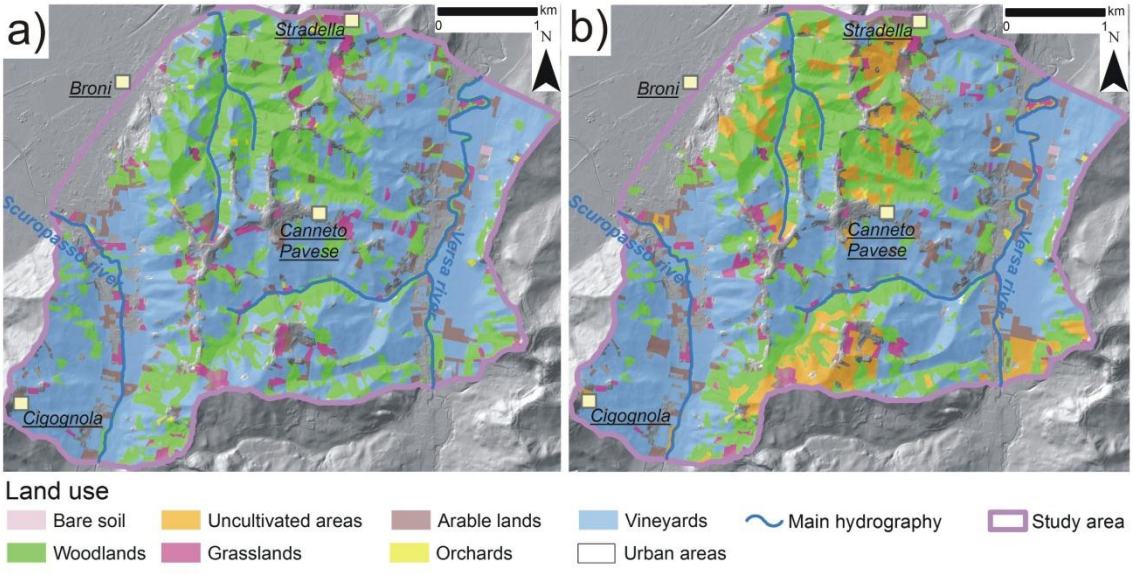


Figure 5: Potential land use scenarios used for the assessment of road susceptibility to shallow landslides in the study area, together with 1980 land use distribution (Fig. 2c): a) Scenario 3, correspondent to the transformation of actual uncultivated areas in woodlands; b) Scenario 4, correspondent to an increase in abandoned areas similar to that one in 1980-2015 period.

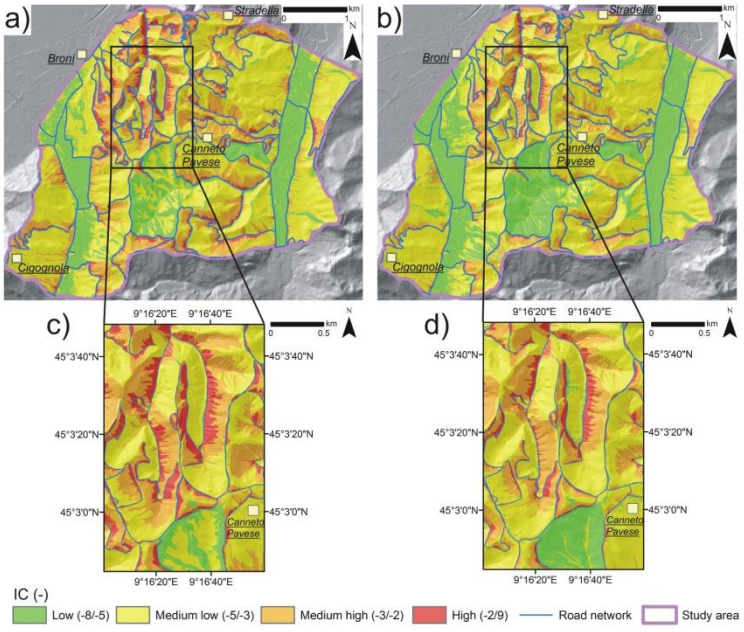


Figure 6: Actual (2015) IC maps corresponding to the linear calculation of the index since W_{lin} (a) and to the non-linear calculation of the index since W_{nl} (b). A detail of the northern sector of the study areas is reported for IC_{lin} (c) and IC_{nl} (d) maps.

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-457
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 24 January 2018

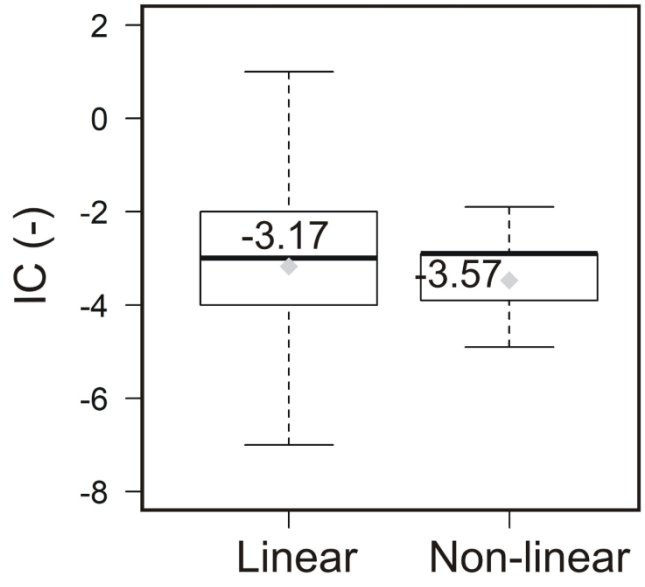


Figure 7: Boxplot of IC values distribution for the actual scenario (2015), for the linear and non-linear calculation of the index.

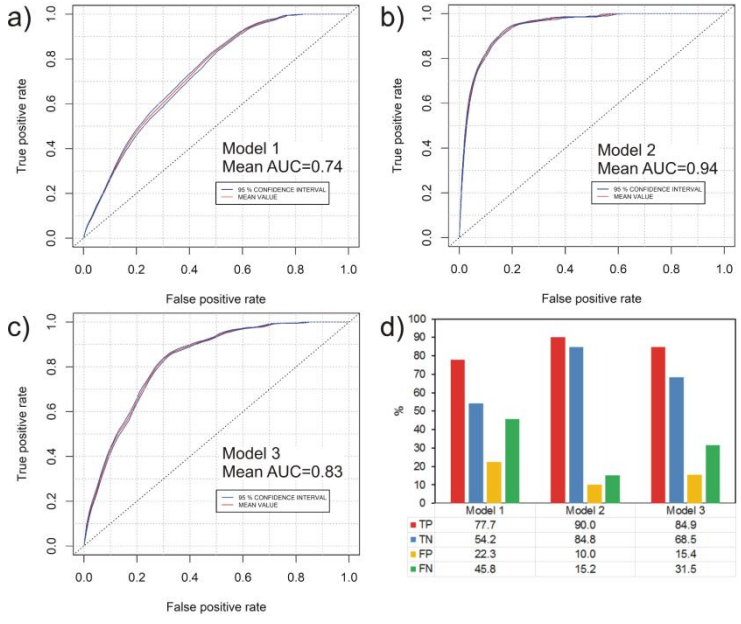


Figure 8: 95% bootstrap confidence bands of ROCs: a) Model 1; b) Model 2; c) Model 3. d) Percentage of true positives (TP), true negatives (TN), false positives (FP), false negatives (FN) of the different models.

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-457
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 24 January 2018

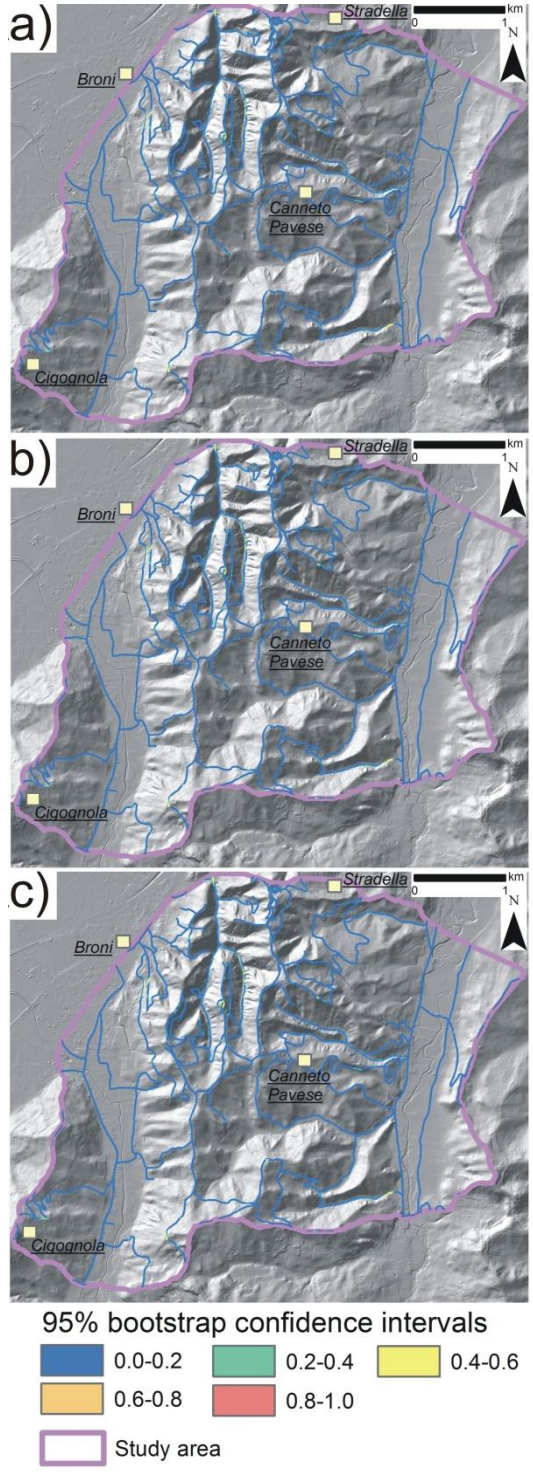


Figure 9: Maps of the amplitude of 95% bootstrap confidence intervals of the probability associated to each pixel of the studied area: a) Model 1; b) Model 2; c) Model 3.

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-457
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 24 January 2018

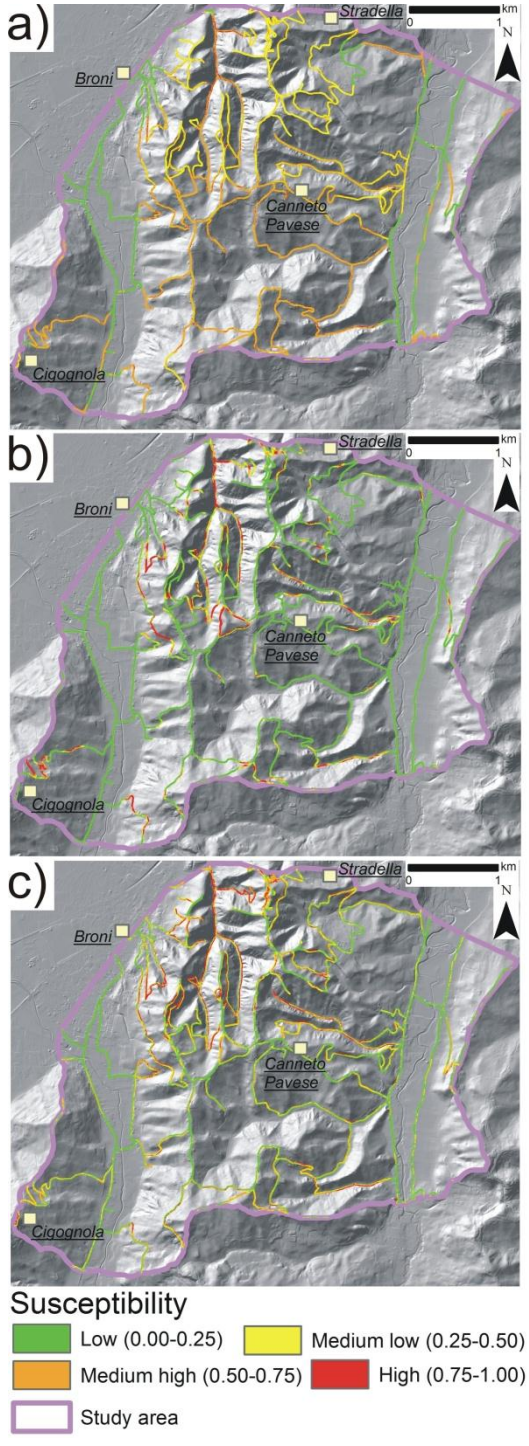
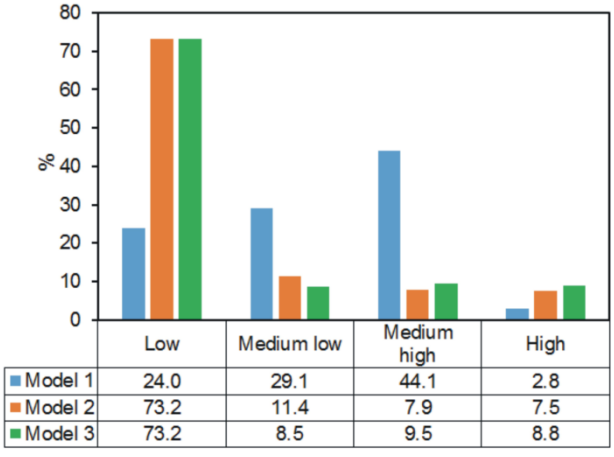


Figure 10: Maps of the susceptibility of the road segments to be affected by shallow landslides: a) Model 1; b) Model 2; c) Model 3.

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-457
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 24 January 2018

Natural Hazards and Earth System Sciences
Discussions



	Low	Medium low	Medium high	High
Model 1	24.0	29.1	44.1	2.8
Model 2	73.2	11.4	7.9	7.5
Model 3	73.2	8.5	9.5	8.8

Figure 11: Percentage of the road network classified with low, medium-low, medium-high or high susceptibility to be affected by shallow landslides for each GAM model.

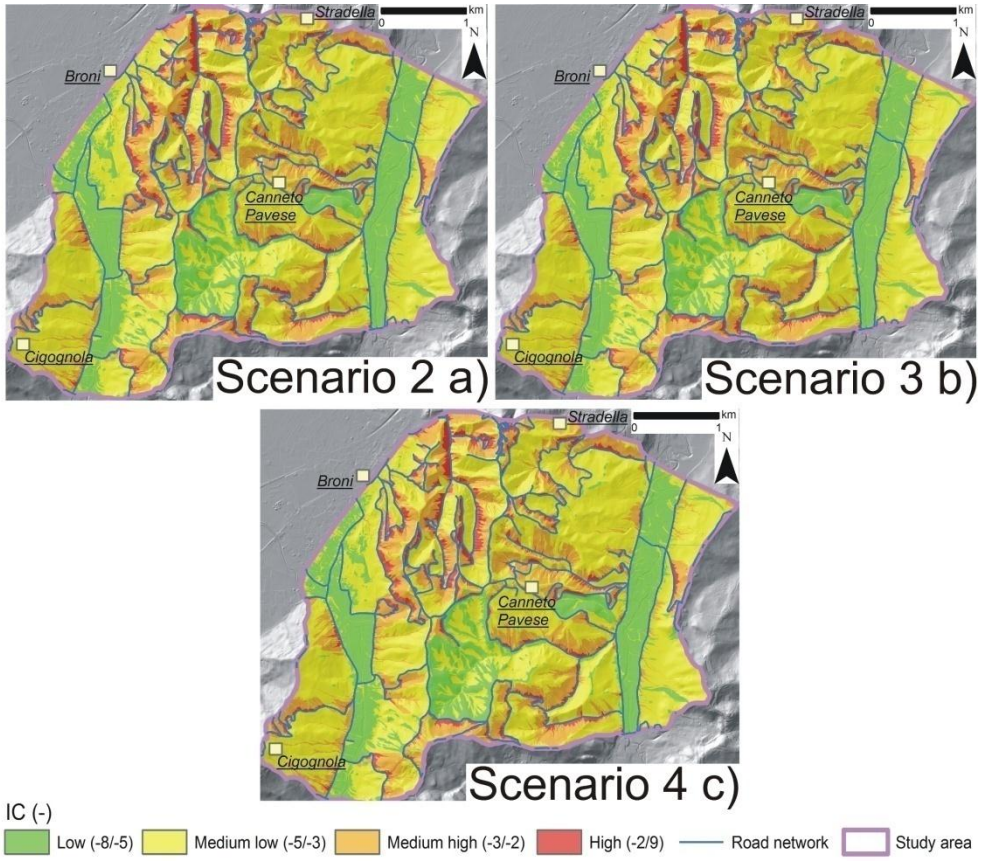


Figure 12: IC_{lin} maps of the different land use scenarios considered: a) Scenario 2: land use distribution equal to that one of 1980 (highest extension of vineyards); b) Scenario 3: transformation of actual uncultivated areas in woodlands; c) Scenario 4: increase in abandoned areas similar to that one occurred in 1980-2015 period.

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-457
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 24 January 2018

Natural Hazards and Earth System Sciences
Discussions

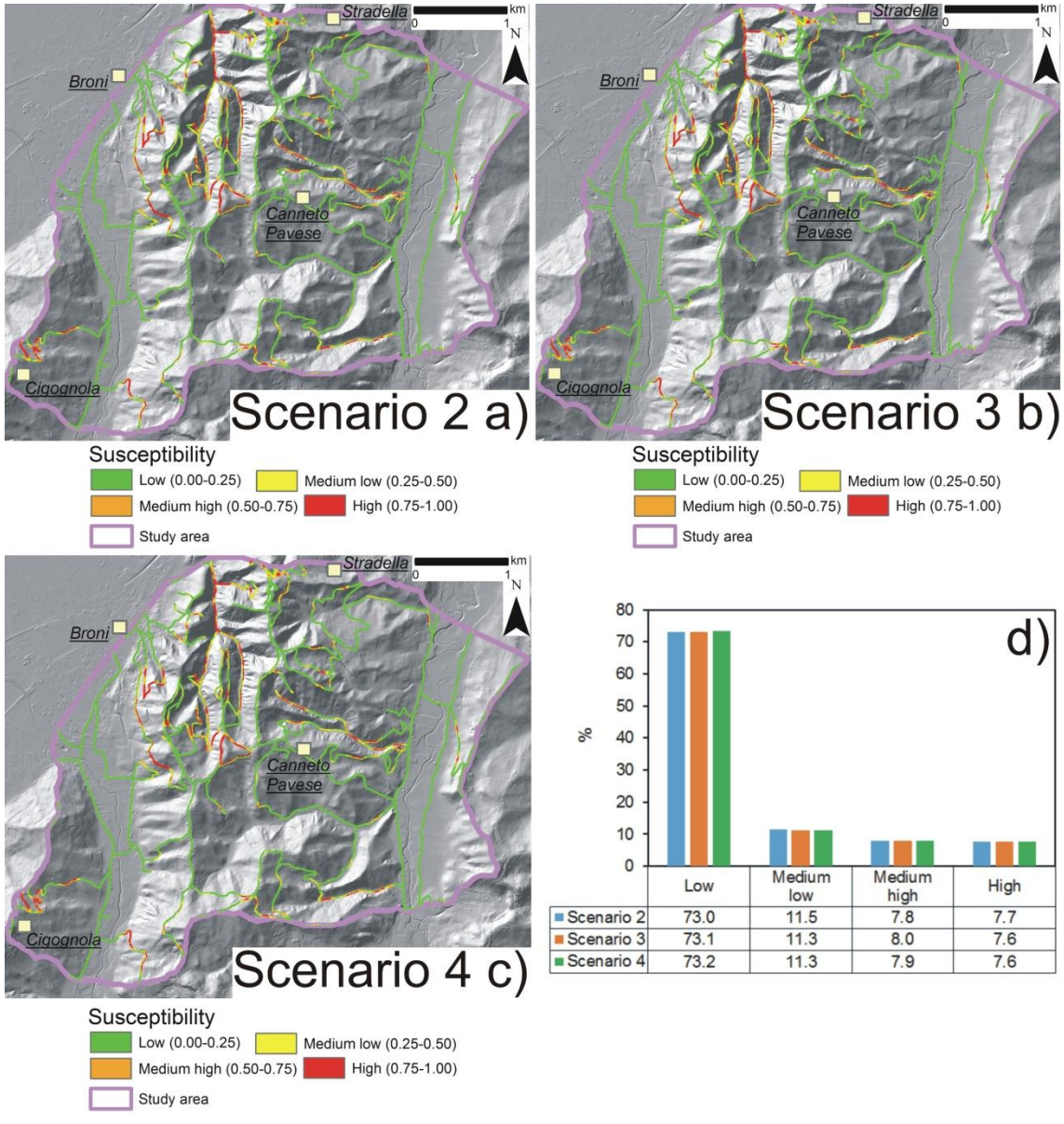


Figure 13: Maps of the susceptibility of the road segments to be affected by shallow landslides according to the different land use scenarios: a) Scenario 2: land use distribution equal to that one of 1980 (highest extension of vineyards); b) Scenario 3: transformation of actual uncultivated areas in woodlands; c) Scenario 4: increase in abandoned areas similar to that one occurred in 1980-2015 period; d) percentage of the road network traits of different susceptibility classes for the considered scenarios.

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-457
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 24 January 2018

Natural Hazards and Earth System Sciences
Discussions

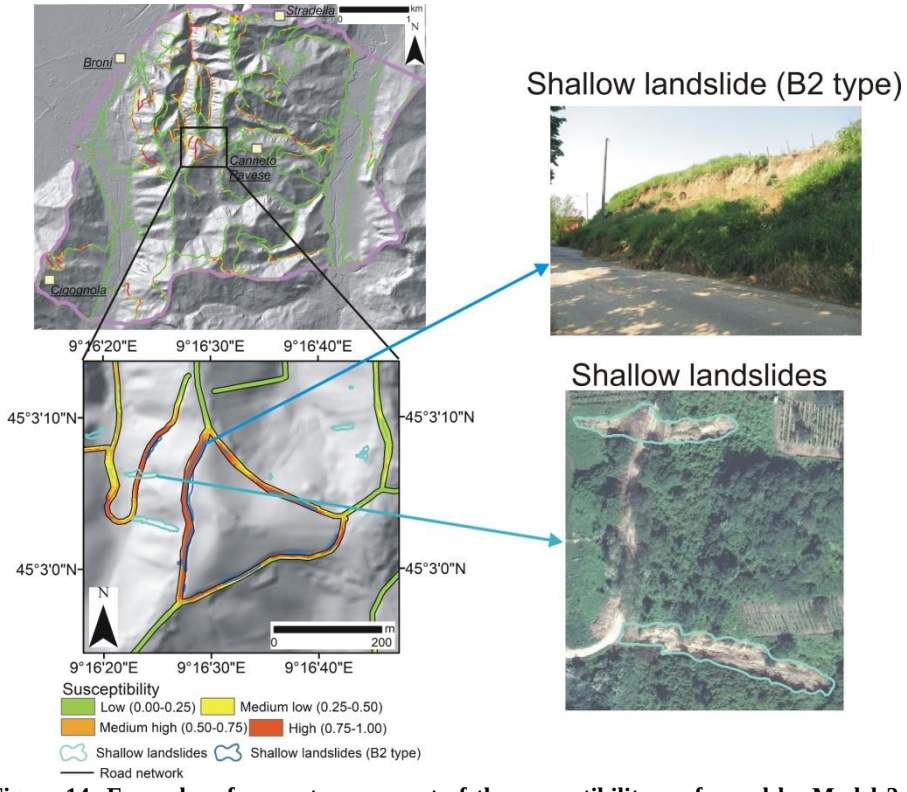


Figure 14: Examples of correct assessment of the susceptibility performed by Model 2, for road sectors hit by B2 type shallow landslides and by other types of phenomena.

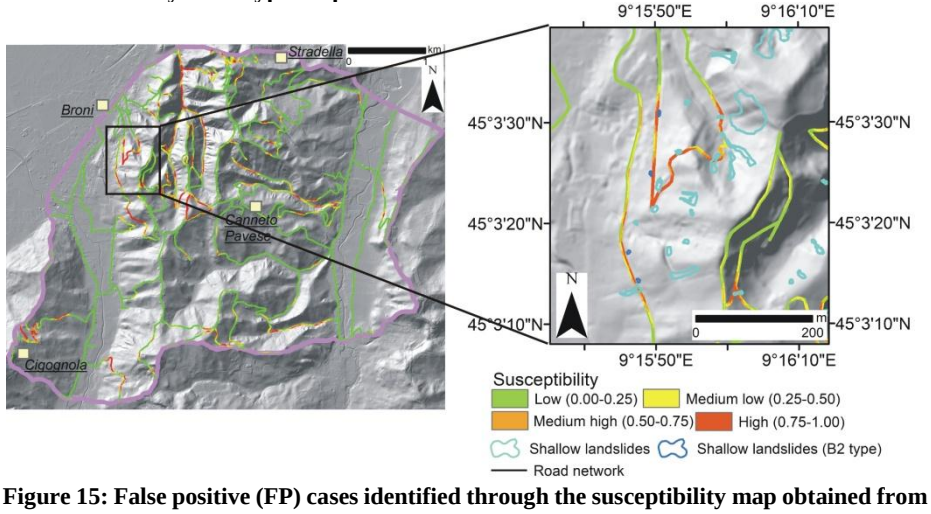


Figure 15: False positive (FP) cases identified through the susceptibility map obtained from Model 2. It is worth noting that FP cases are mostly located close (in a range lower than 250 m) to road sectors already affected by shallow landslides, in similar morphological and connectivity settings.