

## Shallow landslides susceptibility assessment in different environments

Maria Giuseppina Persichillo, Massimiliano Bordoni, Claudia Meisina, Carlotta Bartelletti, Michele Barsanti, Roberto Giannecchini, Giacomo D'Amato Avanzi, Yuri Galanti, Andrea Cevasco, Pierluigi Brandolini & Jorge Pedro Galve

To cite this article: Maria Giuseppina Persichillo, Massimiliano Bordoni, Claudia Meisina, Carlotta Bartelletti, Michele Barsanti, Roberto Giannecchini, Giacomo D'Amato Avanzi, Yuri Galanti, Andrea Cevasco, Pierluigi Brandolini & Jorge Pedro Galve (2017) Shallow landslides susceptibility assessment in different environments, *Geomatics, Natural Hazards and Risk*, 8:2, 748-771, DOI: [10.1080/19475705.2016.1265011](https://doi.org/10.1080/19475705.2016.1265011)

To link to this article: <https://doi.org/10.1080/19475705.2016.1265011>



© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 15 Dec 2016.



Submit your article to this journal [↗](#)



Article views: 459



View related articles [↗](#)





View Crossmark data [↗](#)



Citing articles: 8 View citing articles [↗](#)

## Shallow landslides susceptibility assessment in different environments

Maria Giuseppina Persichillo <sup>a</sup>, Massimiliano Bordoni<sup>a</sup>, Claudia Meisina <sup>a</sup>, Carlotta Bartelletti<sup>b</sup>, Michele Barsanti<sup>c</sup>, Roberto Giannecchini<sup>b</sup>, Giacomo D'Amato Avanzi<sup>b</sup>, Yuri Galanti<sup>b</sup>, Andrea Cevasco<sup>d</sup>, Pierluigi Brandolini<sup>d</sup> and Jorge Pedro Galve<sup>e</sup>

<sup>a</sup>Department of Earth and Environmental Sciences, University of Pavia, Pavia, Italy; <sup>b</sup>Department of Earth Sciences, University of Pisa, Pisa, Italy; <sup>c</sup>Department of Civil and Industrial Engineering, University of Pisa, Pisa, Italy; <sup>d</sup>Department of Earth, Environment and Life Sciences, University of Genova, Genova, Italy; <sup>e</sup>Department of Geodynamics, University of Granada, Granada, Spain

### ABSTRACT

The spatial distribution of shallow landslides is strongly influenced by different climatic conditions and environmental settings. This makes difficult the implementation of an exhaustive monitoring technique for correctly assessing the landslide susceptibility in different environmental contexts.

In this work, a unique methodological strategy, based on the statistical implementation of the generalized additive model (GAM), was performed. This method was used to investigate the shallow landslide predisposition of four sites with different geological, geomorphological and land-use characteristics: the Rio Frate and the Versa catchments (Southern Lombardy) and the Vernazza and the Pogliaschina catchments (Eastern Liguria).

A good predictive overall accuracy was evaluated computing by the area under the ROC curve (AUROC), with values ranging from 0.76 to 0.82 and estimating the mean accuracy of the model (0.70–0.75).

The method showed a high flexibility, which led to a good identification of the most significant predisposing factors for shallow landslide occurrence in the different investigated areas. In particular, detailed susceptibility maps were obtained, allowing to identify the shallow landslide prone areas. This methodology combined with the use of the rainfall thresholds for triggering shallow landslides may provide an innovative tool useful for the improvement of spatial planning and early warning systems.

### ARTICLE HISTORY

Received 22 February 2016  
Accepted 12 November 2016

### KEYWORDS

Shallow landslide susceptibility; generalized additive models; cross-validation and bootstrap techniques; Eastern Liguria; Southern Lombardy

## 1. Introduction

Shallow landslides are phenomena triggered by rainfall events with short duration and high intensity or with long duration and medium–low intensity (Caine 1980; Giannecchini 2006; Giannecchini et al. 2012, 2016). These phenomena affect soil of small thickness (generally less than 2 m) originating from the weathering of the bedrock (residual) and downslope transportation (colluvial). The rainfall-induced shallow landslides are usually characterized by the absence of incipient movement evidence and by high velocity of propagation, especially when the sliding soil portion flows like a fluid down the slope surface, reaching velocities greater than 9 m/s (Campus et al. 1998; Montrasio & Valentino 2007). Moreover, despite the small soil volume involved, the shallow landslides can be

densely distributed across territories, resulting particularly destructive when they evolve into debris flows (Hungri et al. 2001, 2014; Sidle & Ochiai 2006; Crosta et al. 2012).

Due to the absence of warning signs from unstable landslide-prone areas, shallow landslides are very difficult to monitor. As a result, the most effective approach to prevent such phenomena is the use of tools for the spatial and temporal prediction of areas mostly prone to landslides (van Westen et al. 2008; Corominas et al. 2014). The evaluation of shallow landslide hazard, defined as the probability of occurrence of a potentially damaging landslide phenomenon within a given area and in a given period of time (Varnes 1984), is described by three concepts: magnitude, spatial location and time recurrence. The first term refers to the intensity of the natural phenomenon which has influence on its behaviour and destructive power. The second implies the capability to identify the area with the predisposing conditions where the phenomena may occur. The third refers to the temporal frequency of the event (Guzzetti et al. 1999). In particular, the quantitative hazard assessment is based on the landslide magnitude–frequency relation. Without the correct assessment of the expected annual frequency of landslide events of a given magnitude, a quantitative assessment of landslide hazard is not possible. In this case, the problem can only be dealt in terms of landslide susceptibility (Brabb 1984; Guzzetti et al. 2006; Corominas et al. 2014).

The landslide susceptibility gives information on the proneness to landsliding, in terms of initiation areas, on the basis of a set of relevant environmental characteristics. In particular, the main data layers required for landslide susceptibility assessment are landslide inventory data, predisposing and triggering factors (Soeters & van Westen 1996; van Westen et al. 2008; Corominas et al. 2014). Among these data layers, the landslide inventory is the most important, as it gives insight into location of past landslide occurrence, as well as their failure mechanisms.

Therefore, landslide susceptibility maps represent a powerful tool since they provide coherent information on potentially unstable slopes. However, the frequency or the time of occurrence of future landslides are not assessed (Fell et al. 2008). In this context, the landslide susceptibility assessment can be considered the initial step towards the landslide hazard and risk assessment and the end product in land-use planning and environmental impact assessment, as well as important tool in early warning system techniques (Corominas et al. 2014).

The procedures available for the landslide susceptibility assessment can be grouped as based on knowledge-driven, data-driven or physically based methods. The first two methodologies are commonly used in regional hazard analyses, while the last one is used in site specific studies in which the safety factor of the slope is computed (Baeza & Corominas 2001; Corominas et al. 2014). Data-driven methods are quantitative approaches based on the principle that landslides are most likely to occur under similar ground conditions to previous events (Varnes 1984).

Some of the most important advantages of these methods are their objectivity and their easy application at regional scale. Moreover, they allow to manage a large number of predisposing factors simply derived from the elaboration of Digital Elevation Model (DEM), providing automatic and fast procedures in case of lack of exhaustive data.

Among the data-driven approaches (Dai & Lee 2002; Chen & Wang 2007), multivariate analysis is one of the most sophisticated techniques for landslide susceptibility assessment.

In such methodology, slope failure is considered the result of the interaction of several environmental factors that can vary in space and time. In the multivariate statistical models, the spatial distribution of landslides is predicted through the estimation of the relation between the independent predisposing factors and response variables, information of previous landslide occurrence, namely landslide inventory maps (Baeza & Corominas 2001). Multivariate analysis allows the estimation of the relative weight of each contributing factor by means of statistical procedures such as multiple regression or discriminant analysis.

An important statistical improvement was the advance in regression analysis, provided by generalized linear model (GLM) and generalized additive model (GAM; Hastie & Tibshirani 1990). The GLM is a well-established tool for landslide susceptibility modelling and represents a flexible generalization of ordinary linear regression analysis (Brenning et al. 2015).

However, linearity is unrealistic in many environmental modelling situations (Brenning et al. 2015). In the case of geomorphic processes related to landslides occurrence, this nonlinearity can be observed when a hillslope became unstable, as a consequence of changing in multiple hydrological, geomorphological and climatic conditions, which lead to a progressive transformation and movement of slope material (Goetz et al. 2011). Moreover, this nonlinearity opens up a variety of possibilities for complex behaviours that are not possible in linear systems (Phillips 2003, 2006), so the use of simple linear approaches may limit the predictive performance of the models.

For this reason, nonlinear regression techniques, such as the GAM, have recently been applied to landslide susceptibility modelling and hazard zonation (Brenning 2008; Jia et al. 2008; Goetz et al. 2011; Petschko et al. 2014; Brenning et al. 2015). In particular, these new researches highlighted the better predictive performance obtained with the use of the GAM rather than with the GLM (Goetz et al. 2015). This difference was especially related to the inability of the GLM to represent nonlinear relationships, leading to important local bias in predicted landslide susceptibility. This result confirms the GAM ability to describe the complex nonlinear processes characterizing the relationship between landslide occurrence and the independent predisposing factors.

Starting from these important advancements, the aims of this work are as follows:

- (1) the development of a unique methodology strategy, based on the use of the GAM, characterized by simplicity, reproducibility and good predictive efficiency in order to apply it in different environmental contexts;
- (2) the identification of the most influent shallow landslide predisposing factors in different environmental settings.

The GAM was applied to four study sites recently affected by widespread rainfall-induced shallow landslides: the Rio Frate and the Versa catchments (Oltrepò Pavese, Southern Lombardy), the Pogliaschina catchment (Vara Valley, Eastern Liguria) and the Vernazza catchment (Cinque Terre, Eastern Liguria). These sites are characterized by different geological, geomorphological and land-use settings.

## 2. Study areas

In order to develop a unique methodology strategy, applicable to different environmental settings, four catchments were selected. The investigated areas are localized in the north-western Apennines (Figure 1). The Rio Frate (1.9 km<sup>2</sup> wide) and the Versa (38.0 km<sup>2</sup>) catchments are situated in the Oltrepò Pavese (Southern Lombardy); while the Pogliaschina (25.0 km<sup>2</sup>) and the Vernazza (5.7 km<sup>2</sup>) catchments are located in Eastern Liguria in the Vara Valley and along the Tyrrhenian side, respectively.

The Rio Frate, the Vernazza and the Pogliaschina catchments are characterized by steep slopes and hard rock units, namely arenaceous conglomeratic bedrock, sandstone–siltstone flysch, sandstone–claystone flysch and pelitic complex, respectively. These areas were affected by different shallow landslide typologies. In particular, the Rio Frate catchment was mostly affected by roto-translational slides evolving into earth flows, while the Pogliaschina and Vernazza catchments were completely affected by complex, translational debris slide-flows and debris avalanches. Moreover, the Rio Frate and the Vernazza catchments were characterized by a high percentage of abandoned lands previously cultivated with vineyards and terraced vineyards/olive groves, respectively; on the contrary, the Pogliaschina catchment was instead characterized by a slight reduction of agricultural areas replaced by woodlands in the last 50 years.

The Versa catchment is mostly composed of flysch with clay component, relatively uniform slopes with medium–low topographic gradient and characterized by intensively cultivated areas, especially vineyards. This area was mainly affected by roto-translational slides that evolved into earth flows.

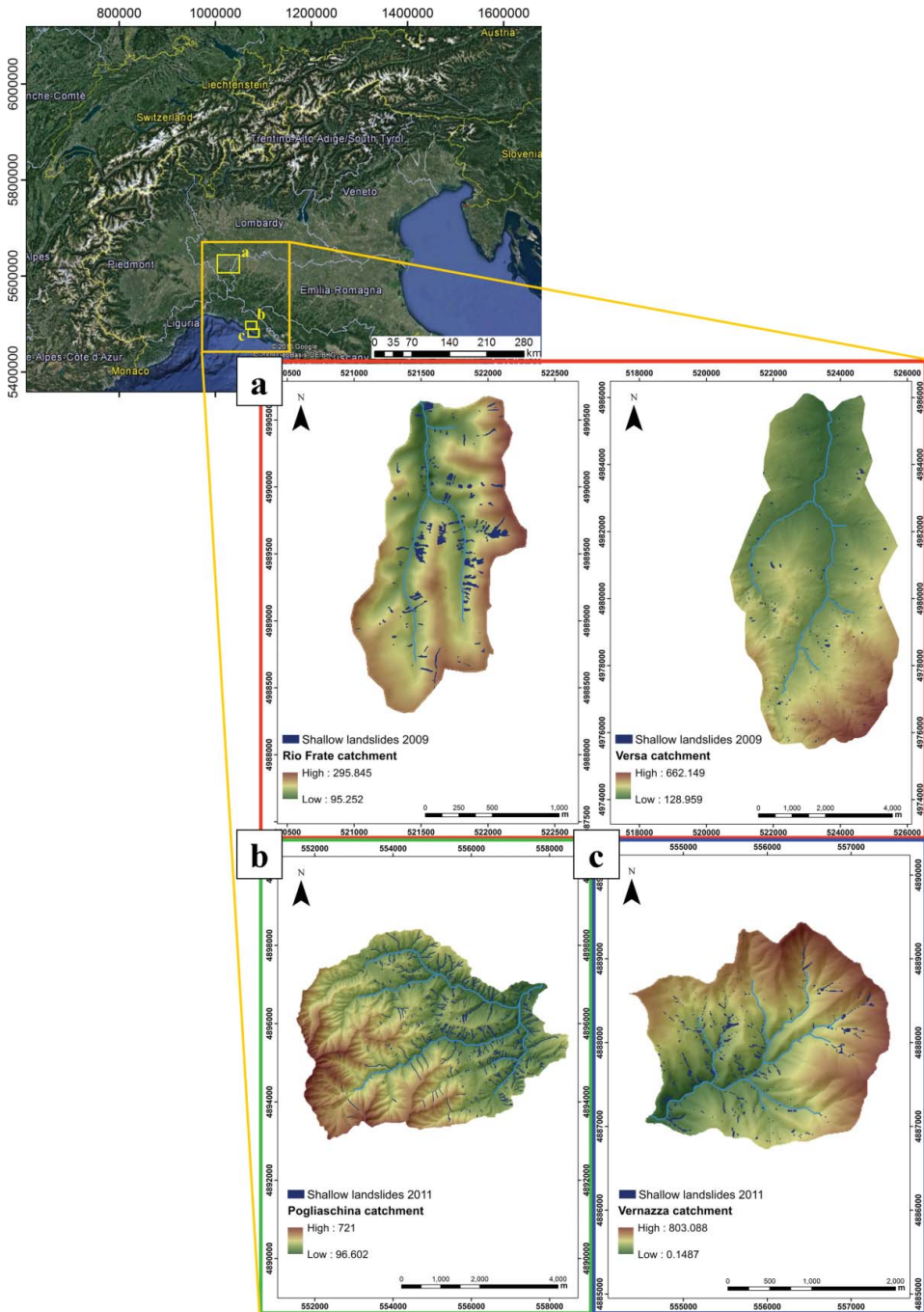


Figure 1. (a) Upper left the Rio Frate catchment; Upper right the Versa catchment; (b) Pogliaschina catchment; (c) Vernazza catchment.

The four study catchments were recently affected by widespread rainfall-induced shallow landslides occurred on 27-28 April 2009 (Rio Frate and Versa catchments) and on 25 October 2011 (the Vernazza and Pogliaschina catchments). The 27-28 April 2009 and the 25 October 2011 landslide inventories were used as response variables in the landslide susceptibility assessment.

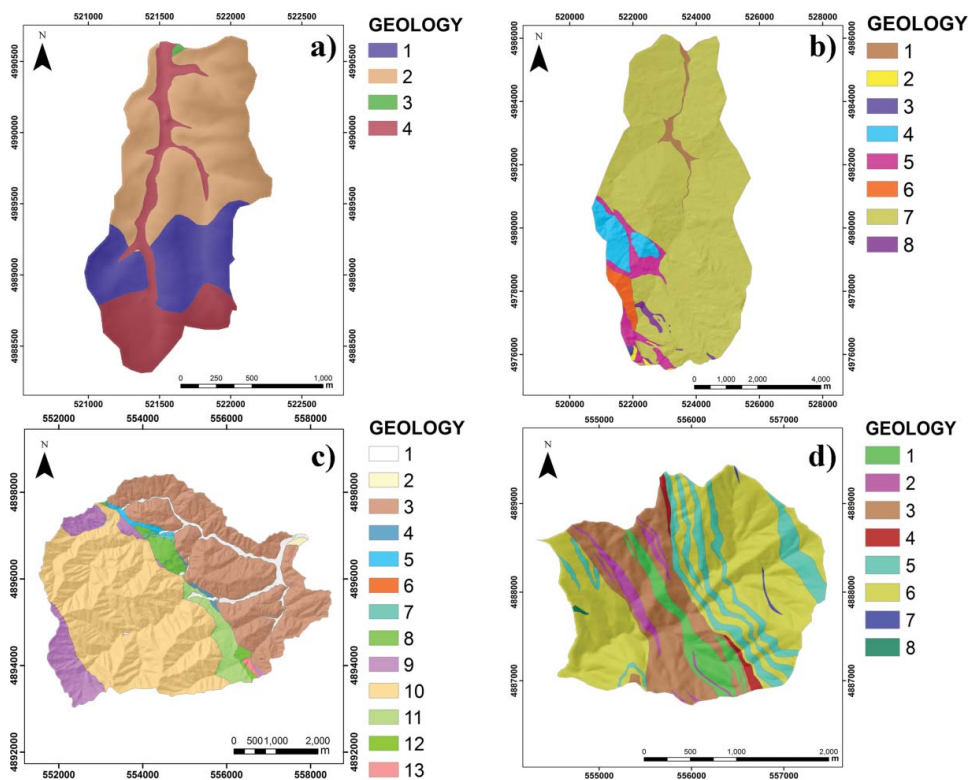
## 2.1. Rio Frate catchment

The Rio Frate catchment (Figure 2(a)) has an elevation ranging between 95 and 295 m a.s.l. The slopes have medium–high gradient, which can reach values higher than  $35^\circ$ , and they end in narrow small valleys, ranging between 55 and 348 m<sup>2</sup>.

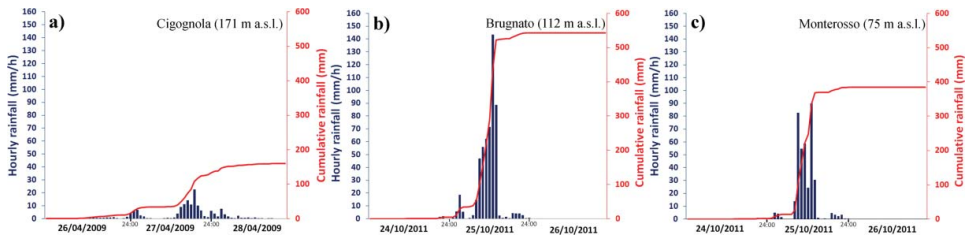
In this area the bedrock is characterized by a Mio-Pliocene succession named ‘Serie del Margine’ (Vercesi & Scagni 1984). In particular, medium low-permeable arenaceous conglomeratic bedrock (Monte Arzolo Sandstones, Rocca Ticozzi Conglomerates) overlies impermeable silty-sandy marly bedrock and evaporitic chalky marls and gypsum (Sant’Agata Fossili Marls, Gessoso-Solfifera Formation). The superficial soils, derived by bedrock weathering, are mainly clayey sandy silts and clayey silty sands. The soil thickness ranges from few centimetres to less than 2 m.

The Rio Frate catchment was affected by a significant land-use change in the last 60 years. Since 1954 the area has been characterized by increasing of abandoned lands, previously cultivated by vineyards. The current land use consists of vineyards (26%) and woodlands (44%), mostly derived from the abandonment of vineyards especially occurred after the 1980s.

This area was affected by hundreds of shallow landslides during the 27–28 April 2009 event (Zizioli et al. 2013). This event was characterized by 160 mm of cumulated rainfall in 62 hours (representing 20% and 25% of the mean annual rainfall of 1921–1979 and 2004–2013 period,



**Figure 2.** Geology maps: (a) Rio Frate catchment: 1. Monte Arzolo Sandstones; 2. Rocca Ticozzi conglomerates; 3. Gessoso-Solfifera Fm.; 4. Sant’Agata Fossili marls; (b) Versa catchment: 1. Alluvial deposits; 2. Ranzano sandstones; 3. Ranzano sandstones (arenaceous lithofaces); 4. Ranzano sandstones (marly lithofaces); 5. Monte Piano marls; 6. Val Luretta Fm. (arenaceous lithofaces); 7. Val Luretta Fm. (calcareous lithofaces); 8. Varicolori clays; (c) Pogliaschina catchment: 1. Alluvial deposits (current); 2. Alluvial deposits (recent); 3. Monte Gottero sandstones Fm.; 4. Val Lavagna Fm.; 5. Argille a Palombini Fm.; 6. Diaspri di Monte Alpe Fm.; 7. Gabbri Fm.; 8. Serpentiniti Fm.; 9. Canetolo clays and limestones Fm.; 10. Macigno Fm.; 11. Scaglia Toscana Fm.; 12. Maiolica Fm.; 13. Diaspri Fm.; (d) Verazzza catchment: 1. Ponte Bratica sandstones; 2. Groppo del Vescovo limestones; 3. Canetolo clays and limestones; 4. T. Pignone marls; 5. Macigno Fm.; 6. Macigno Fm. (sandstones lithofaces); 7. Macigno Fm. (silty-pelitic lithofaces); 8. Macigno Fm. (silty-marl lithofaces).



**Figure 3.** Shallow landslide triggering rainfall: (a) event occurred on 27–28 April 2009 in the Rio Frate and the Versa catchments and recorded by the Cigognola rain gauge; (b) event occurred on 25 October 2011 in the Pogliaschina catchment and recorded by the Brugnato rain gauge; (c) event occurred on 25 October 2011 in Vernazza catchment and recorded by the Monterosso rain gauge.

respectively) with a maximum intensity of  $22 \text{ mm h}^{-1}$  (from 18:00 to 19:00 UTC) on 27 April 2009 (Zizioli et al. 2013; Figure 3(a)).

245 shallow landslides were mapped. These were characterized by an average length of about 35 m and their extent varied from 13 to  $6289 \text{ m}^2$ , with an average of about  $473 \text{ m}^2$ . The slide surface depth ranges between 0.90 m and 1 m (Table 1).

The shallow landslides were ranked according to Cruden and Varnes (1996) classification. Most of them were classified as roto-translational slides evolving into flows, with width/length ratio  $> 1$ .

## 2.2. Versa catchment

The Versa catchment (Figure 2(b)) covers an area of about  $38 \text{ km}^2$ , with altitude ranging between 128 and 662 m a.s.l. The slopes have a low–medium gradient, with values commonly ranging between  $15$  and  $25^\circ$ . It has a predominant clayey-marly component, constituted by flysch deposits (Ranzano Sandstones, Monte Piano Marls, Val Luretta Formation). Soil slope covers have a clayey texture and their thickness can reach values higher than 3–4 m.

In the Versa catchment an important modification of agricultural practices occurred after 1954, when sowed areas were substituted by vineyards, which currently represent the most widespread cultural practice (65%).

196 shallow landslides were triggered during the 27–28 April 2009 event. The triggering event was the same occurred in the Rio Frate catchment (Figure 3(a)). The shallow landslides were characterized by an average length of 48 m and extent ranging between 6 and  $8098 \text{ m}^2$ , with an average extension of  $491 \text{ m}^2$ . The slide surface depth varied between 0.90 and 1 m (Table 1). Also in this case, the shallow landslides were ranked according to Cruden and Varnes (1996) classification and most of them were classified as roto-translational slides evolving into flows, with width/length ratio  $> 1$ .

## 2.3. Pogliaschina catchment

The Pogliaschina catchment is located in the Vara Valley (Northern Apennines; Figure 1). It is  $25 \text{ km}^2$  wide and has a maximum altitude of about 720 m a.s.l., while the valley bottom is about

**Table 1.** The main morphometric characteristics of the landslide inventories of each study areas.

Catchment	N° landslides	Min area ( $\text{m}^2$ )	Max Area ( $\text{m}^2$ )	Average area ( $\text{m}^2$ )	Average length (m)	Slide surface depth (m)
Rio Frate	245	13	6289	473	35	0.90–1
Versa	196	6	8098	491	48	0.90–1
Pogliaschina	658	50	18502	1190	95	1–2
Vernazza	473	5.5	6307	322	74	1.5 (on woodlands) 2.5 (on terraced slopes)

95–100 m a.s.l. The slopes are characterized by high gradient, since they can reach values higher than  $45^\circ$ . The bedrock is mainly composed of medium and coarse quartz-feldspathic sandstone turbidite (Macigno Fm.; Tuscan *Nappe* Unit) and quartz-feldspathic greywacke sandstone-siltstone turbidite (Arenarie di Monte Gottero Fm.; Gottero Unit) and is covered by a soil thickness ranging from 0.5 to 1.5 m. From a tectonic point of view, the Vara Valley occupies a depression separated from La Spezia Dorsal to the west and from M. Picchiara-M. Cornoviglio Dorsal to the east, that originated from the combination of two main tectonic phases associated with the Northern Apennine chain formation (Raggi 1985; Figure 2(c)).

The Vegetation cover is mainly represented by woodland, characterized by hardwood (mostly chestnut trees), coniferous (principally maritime pines) and mixed hardwood and coniferous forests (93% of the whole basin). Vineyards, olive groves and other plantations occupy about 6% of the area (D'Amato Avanzi et al. 2014). Since 1960 these agricultural areas have been affected by a small reduction (9%) and further substituted by woodlands.

On 25 October 2011, a cumulated rainfall of 538.2 mm in 24 h was recorded by the Brugnato rain gauge, with a rainfall intensity of  $143.4 \text{ mm h}^{-1}$  (from 13:00 to 14:00 UTC) and up to  $108.1 \text{ mm h}^{-1}$  in 6 h (from 9:00 to 15:00 UTC; D'Amato Avanzi et al. 2014; Figure 3(b)). A total of 658 shallow landslides were mapped, 569 of which were classified as complex, translational debris slide-flows (Cruden & Varnes 1996). They were usually superficial (0.3–1.5 m thick), linear (width/length ratio 0.03–0.5) and involved mostly coarse-grained soil and sometimes portions of the fractured bedrock. 89 landslides were classified as roto-translational slides (Cruden & Varnes 1996), with a small size and a width/length ratio  $> 1$  (Bartelletti et al. 2015). The landslides were characterized by a minimum area of  $50 \text{ m}^2$  and a maximum area of  $18502 \text{ m}^2$ , with an average of about  $1190 \text{ m}^2$ . The average length of the landslides was approximately 95 m and their slide surface was localized at depth of about 1–2 m (Table 1).

#### 2.4. Vernazza catchment

The Vernazza catchment (Eastern Liguria; La Spezia province) is located along the Tyrrhenian side of the northern Apennines (Figure 1). It is part of the Cinque Terre area, which was declared World Heritage Site by Unesco and it is included in the Cinque Terre National Park.

The bedrock is mainly composed of a sandstone–claystone flysch (Macigno Fm., Tuscan *Nappe*) and a pelitic complex (Canetolo Shales and Limestones, Canetolo Unit), which are part of a wide overturned antiform megafold whose axes strikes  $150^\circ\text{N}$  (Regione Liguria 2006; Figure 2(d)).

It shows typical geomorphological features characteristic of most of the Ligurian coastal catchments: small area (about  $5.7 \text{ km}^2$ ), very steep slopes due to the proximity of mountains to the sea and short streams, often controlled by tectonics, with considerable erosive power and capacity to transport sediment because of their steep profiles. More than 50% of the terrain gradient ranges between  $30^\circ$  and  $40^\circ$ .

Unusual land-use conditions characterize the slopes within the Vernazza catchment (Cevasco et al. 2013a; 2014). Approximately 50% of these slopes, located in the middle and lower part of the catchment, were terraced during the past millennium for vineyards and olive groves. Following the exodus of farmers in the last century, terraced slopes have been progressively abandoned, leading to increasing geomorphological instability (Terranova et al. 2006; Cevasco et al. 2015; Brandolini 2016). Currently, only approximately 20% of the terraced areas are still cultivated. In abandoned terracing, the dry stone walls retaining the terraces are in poor condition because of lack of maintenance. However, due to their old age, also the dry stone walls retaining terraced areas still cultivated do not guarantee, in most cases, an effective drainage of surface water. The upper part of the basin is characterized by woodlands and scrublands. Due to reworking of debris covers for terracing, the soil thickness is greater on agricultural terraces (up to 2.5 m) than on woodlands (up to 1.5 m).

Whilst large rock landslides and rock failures are widespread on the coast of Liguria (Brandolini et al. 2007; Carobene & Cevasco 2011; De Vita et al. 2012; Corradi et al. 2013), within the Vernazza catchment, due to its morphological features, shallow phenomena triggered by rainfall are prevalent (Cevasco et al. 2013a). On 25 October 2011, high-intensity rainfall affected the Cinque Terre area,



triggering more than 500 shallow landslides in the Vernazza catchment. A rainfall amount of 382 mm was recorded in 24 h by the Monterosso rain gauge located 3 km to northwest of Vernazza. At Monterosso, rainfall intensities of 90 mm h<sup>-1</sup>, 195 mm/3 h and 350 mm/6 h were recorded between 9:00 and 15:00 UTC (A.R.P.A.L.-C.F.M.I.-P.C. 2012; Cevasco et al. 2014; Figure 3(c))

A total of 473 landslides were mapped. They were characterized by an average length of about 74 m. Moreover, their extent is ranging between 6 and 6307 m<sup>2</sup>, with an average value of 322 m<sup>2</sup> (Table 1). The slide surface depth varies according to the location of the landslides and the characteristics of the soil. In fact, the failures of eluvial and colluvial soils are located at a depth up to 1.5 m on woodland while the failures of artificially reworked deposits are located at a depth up to 2.5 m on terraced slopes (Table 1; Cevasco et al. 2013b, 2014). The shallow landslides recorded were classified following the landslide classifications of Cruden and Varnes (1996) and Hungr et al. (2001). Most of the landslides occurred on 25 October 2011 were classified as debris avalanches (Cevasco et al. 2014).

### 3. Method

The methodology, applied for the shallow landslide susceptibility assessment, is based on the application of a multiple regression model based on GAM (Hastie & Tibshirani 1990).

The procedure for the data analysis is made of two phases (Figure 4):

- (1) Pre-processing of input data in order to extract a list of predictor variables;
- (2) Statistical implementation of GAM.

#### 3.1. Pre-processing of input data

##### 3.1.1. The predictor variables

During the pre-processing phase, 12 predictor variables were identified, according to their influence on shallow landslide initiation (Table 2).

Nine different predisposing factors, representing morphological and hydrological parameters, were extracted by 10-m resolution DEMs, through a set of tools implemented in SAGA GIS (System for Automated Geoscientific Analyses). All DEMs were antecedent to the shallow landslide events.

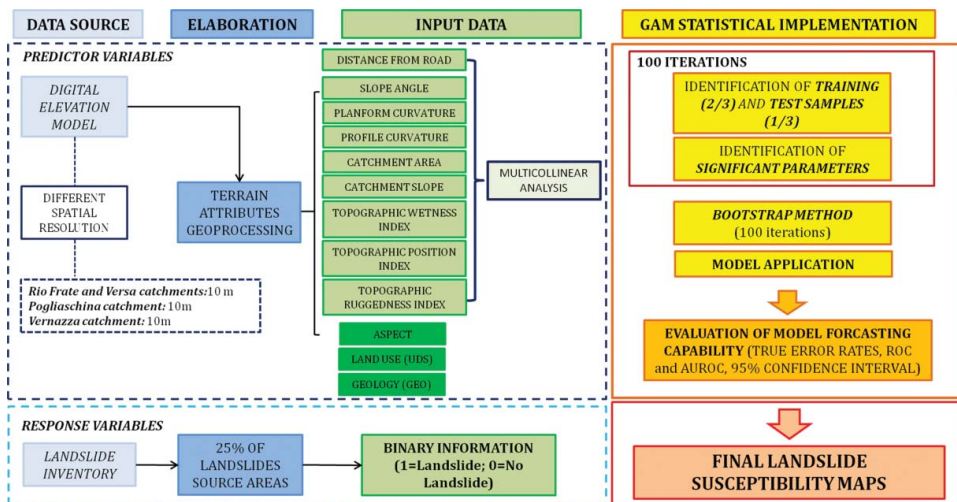


Figure 4. GAM methodology scheme.

**Table 2.** Explanatory variables used in GAM analysis and their influences on shallow landslides triggering.

Variable	Data source	Influence on shallow landslide mechanisms
Slope (SL)	DEM	Affects the magnitude of the downslope driving forces and water distribution (Baeza et al. 2010; Catani et al. 2013)
Aspect (ASP)	DEM	Having influences on solar radiation intensity, which controls the local temperature and evaporation and therefore soil moisture (Van Westen et al. 2008; Demir et al. 2013)
Planform curvature (PLA)	DEM	Divergent or convergent curvature: Influences on the water run-off concentration (Dai et al. 2002; Goetz et al. 2011)
Profile curvature (PRO)	DEM	Concave or convex curvature: Influences on near-surface acceleration or deceleration of flow ; potential erosion or deposition rate (Dai et al. 2002; Goetz et al. 2011)
Catchment area (CA)	DEM	Proxy for soil moisture and soil depth (Brenning et al. 2015)
Catchment slope (CS)	DEM	Influences on the intensity of the destabilizing forces upslope (Brenning et al. 2015)
Topographic wetness index (TWI)	DEM	Influences on the soil moisture content and the groundwater conditions (Seibert et al. 2007)
Topographic position index (TPI)	DEM	Proxy on the effects of the location of a point on a landscape to the shallow landslides occurrence (Weiss 2001)
Terrain ruggedness index (TRI)	DEM	Proxy to quantify the landscape heterogeneities (Riley et al. 1999)
Distance from roads (RD)	DEM	Influence on changing of hillslope geometry and hydrology (Devkota et al. 2013; Demir et al. 2013)
Land use (LU)	Land-use database	Mechanical and hydrological effects of different vegetation types on slope stability (Tasser et al. 2003; Pereira et al. 2012)
Geology (GEO)	Geological database	Influence on the weathering of the bedrock and the resulting soils (Dai and Lee 2002; Catani et al. 2005, 2013; Costanzo et al. 2012)

The DEMs of the Rio Frate and the Versa catchments were obtained by the Regional Technical Map (original scale 1:10000 ed. 1980–1994). The DEM of Pogliaschina derived from the Regional Technical Map (original scale 1:5000 II ed. Regional Topography 3D/DB, 2007) with at 5-m resolution resampled to 10 m. The DEM of Vernazza catchment was realized using LiDAR data acquired in 2008 and 2010 by the Ministry for Environment, Land and Sea (Ministero dell’Ambiente e della Tutela del Territorio e del Mare, MATTM), within the framework of the ‘Extraordinary Plan of Environmental Remote Sensing’ (PST-A).

The geomorphic and hydrological variables derived from DEMs are the following (Table 2): Slope (SL), Aspect (ASP), Planform curvature (PLA), Profile curvature (PRO), Catchment area (CA), Catchment slope (CS), Topographic wetness index (TWI), Topographic position index (TPI), and Terrain ruggedness index (TRI). These variables were selected according to their capacity to outline destabilizing factors (Montgomery & Dietrich 1994; Brenning 2008). In particular, they simplify complex geomorphological relationships allowing to represent several processes such as sub-surface flow convergence, increased soil saturation and shear strength reduction (Goetz et al. 2011).

In addition to the DEM-derived predisposing factors, the Euclidean distance from roads (RD) was calculated. RD was selected as predisposing factor since the landslide inventories were composed also by landslides developed in correspondence of roads. Land use (LU) and Geology (GEO) were also considered as predictors, according to their influence on shallow landslide mechanisms and spatial distribution (Table 2).

SL, ASP, PLA and PRO were calculated based on local polynomial approximations, according to Zevenbergen and Thorne (1987). SL is one of the most important factors, as it strongly controls the shear forces acting on hillslopes and the water distribution (Baeza et al. 2010; Catani et al. 2013). ASP can play a key role in landslide susceptibility, as it may have influence on the exposition of the terrain to different amounts of rainfall and solar radiation, thus conditioning the local temperature and evaporation and, consequently, the soil moisture and the vegetation growth (van Westen et al. 2008; Demir et al. 2013). ASP was transformed from continuous into categorical variable to avoid the presence of ‘No Data’ in correspondence of flat areas. PLA and PRO represent the topographic influence of local morphology on slope hydrology and soil erosion and deposition. In particular,

PLA is the curvature of the surface perpendicular to the direction of the maximum slope, which controls the convergence and divergence of topography and the sub-surface water flow; PRO is the curvature of surface parallel to the direction of the maximum slope, which characterizes the near-surface acceleration or deceleration of flow down the slope, influencing the potential erosion or deposition rate and, consequently, the soil depth. (Dai et al. 2002; Goetz et al. 2011). CA and CS were derived using the multiple-flow direction algorithm (Quinn et al. 1991). CA was then transformed to the natural logarithm in order to reduce skewness and used as proxy for soil moisture and soil depth (Brenning et al. 2015). CS is another important factor which influences the intensity of the destabilizing forces upslope (Brenning et al. 2015). TWI highlights the tendency of water to accumulate at any point in the drainage basin and the tendency of the water to move along a slope by the action of the gravitational forces, so that it can be easily correlated to the soil moisture content and the groundwater conditions (Seibert et al. 2007). TPI provides a simple proxy to study the effects of the location of objects on a landscape. In the case of landslides, this allows to understand their localization in relation to the slope elevation. It was calculated by comparing the elevation of a cell to the mean elevation of the surrounding pixels in a ring buffer of around 1000 m of radius (Weiss 2001). TRI, calculated as the sum of the change in altitude between a cell of the grid and its eight neighbouring cells (Riley et al. 1999), was used to quantify the landscape heterogeneity, which could have effects on the localization of shallow landslide triggering areas.

RD was calculated by using the street network shapefiles available for each study areas. In the Rio Frate, the Versa and the Pogliaschina catchments it consisted only of the main roads, whereas, in the Vernazza catchment it included also the footpaths. This variable has been successfully used in landslide susceptibility assessments as landslide predisposing factor (Demir et al. 2013; Devkota et al. 2013). Effectively the construction of roads, which sometimes implies road cuts or fills and culverts, can lead to anthropogenic modification of the hillslope profile or drainage system.

LU and GEO were derived by means of GIS analyses from specific thematic maps. LU was used to consider the mechanical and hydrological effects of different vegetation types as well as the influence of terracing on slope stability (Tasser et al. 2003; Pereira et al. 2012). GEO allowed to represent the geomechanical and hydraulic properties of the bedrock; it influences the characteristics of the soil coverage (Dai & Lee 2002; Catani et al. 2005, 2013; Costanzo et al. 2012). The evaluation of GEO was carried out considering the different geological formations as reported on the available geological maps.

Moreover, the LU map coming from DUSAF of 2007 (Lombardy Region) was used for the Rio Frate and the Versa catchments. This map was realized using colour and infrared orthophotos IT2007 (made by Blom CGR, pixels 50 cm). In the Pogliaschina catchment, the LU map of 2012 (scale 1:10000; Regione Liguria 2013) was utilized. In the Vernazza catchment a detailed LU map was expressly prepared through analysis of high-resolution aerial photographs and field surveys (Cevasco et al. 2013a). The air photo analysis was carried out on digital georeferenced orthophotos, provided by Liguria Regional Administration, taken by the Air Service of Remote Sensing and Monitoring of Civil Protection of Friuli Venezia Giulia Regional Administration (11 November 2011 flight; ground resolution from 3 cm to 50 cm, according to the altitude). Concerning the GEO variable, the geological maps of the Rio Frate and Versa catchments were realized by the Department of Earth and Environmental Sciences of University of Pavia by means of field surveys. The geological map of the Pogliaschina catchment (scale 1:10000) includes data coming from geological maps performed by Abbate et al. (2005), Bortolotti et al. (2011) and Puccinelli et al. (2015). The geological map of Vernazza catchment derives from the CARG project (original scale 1:25000; Regione Liguria 2006).

### 3.1.2. The response variables

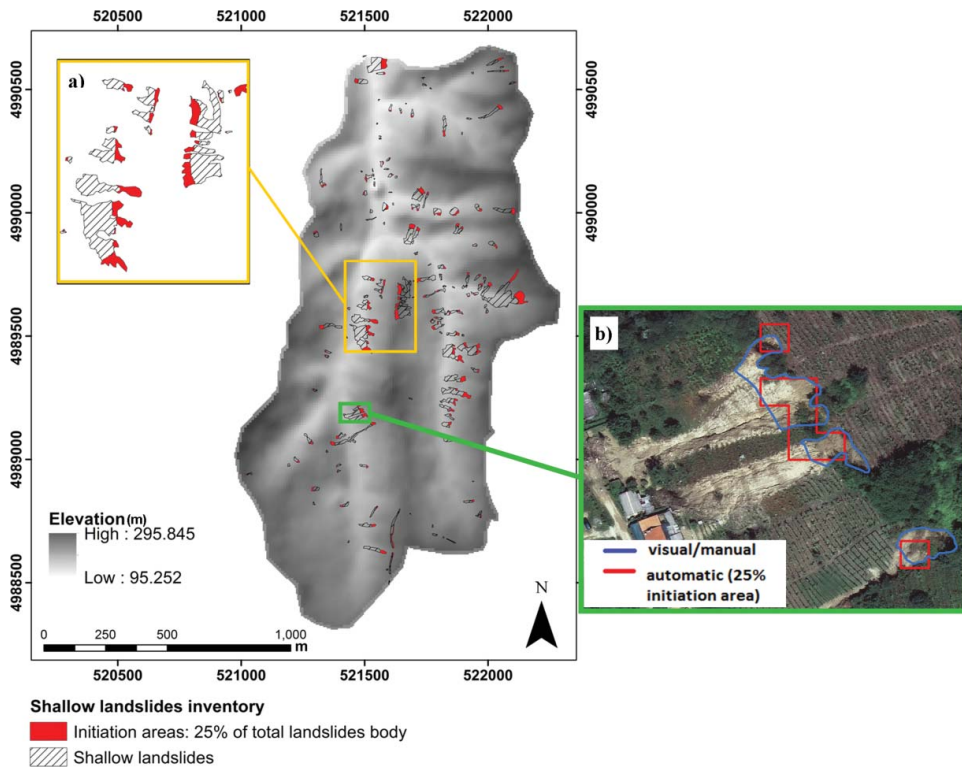
Detailed landslide inventory maps were prepared for all the study areas. The Rio Frate and the Versa inventory maps are referred to the shallow landslide event occurred on 27–28 April 2009. Post-event colour aerial photographs (18 May 2009) at resolution of 15 cm (photo scale of 1:12000) were examined. Detailed field surveys were also conducted to detect slope failures and to identify landslide

features (e.g. scarp and body). Aerial photo interpretation, coupled with field surveys, allowed detecting about 245 and 196 shallow landslides in the Rio Frate and the Versa catchments, respectively.

In the Vernazza and the Pogliaschina catchments the landslide inventories related to the 25 October 2011 event were used. The inventories were obtained by means of the analysis of post-event aerial photographs, digital georeferenced orthophotos (both provided by Liguria Regional Administration) and detailed field surveys. The aerial photos belonged to BLOM C.G.R.S 28 October 2011 flight, while, the orthophotos were the same deriving the LU map (taken by the Air Service of Remote Sensing and Monitoring of Civil Protection of Friuli Venezia Giulia Regional Administration on 28 November 2011). For what concern the Vernazza catchment, we referred to the inventory available from Cevasco et al. (2013a). In the Pogliaschina catchment, the Google Earth images were also used in order to analyse areas not covered by aerial reliefs.

All landslide inventories consist of polygons which bound the whole landslide perimeter. A semi-automatic method, already used in Galve et al. (2015), was implemented to extract each landslide source areas. This procedure allowed to define the source areas by selecting the 25% of the pixels with the highest elevation in each landslide (Figure 5(a)). The extracted landslide source areas were applied in GAM as the binary response variable, assigning 0 or 1 in case of landslide absence or presence, respectively.

The source areas automatically extracted were compared with those selected manually by visual interpretation of orthophotos. This analysis was useful to assess the uncertainty of the automatic procedure. The difference obtained is due to the transformation from vector to raster format at



**Figure 5.** Source areas extracted for the shallow landslides event of 27–28 April 2009 in the Rio Frate catchment: (a) example of 25% of total landslide body defined; (b) comparison between the source areas automatically extracted in raster format and those selected manually by visual interpretation of orthophotos.

10 m resolution. This leads to an overestimation and underestimation of about 10 m due to the raster resolution used (Figure 5(b)).

### 3.2. Statistical implementation of GAM

The second phase of the developed methodology (Figure 4) is referred to the implementation of the GAM (Hastie & Tibshirani 1990).

The GAM represents a semi-parametric extension of the GLM. Its basic assumption is to replace the linear function, used in the GLM, with an empirically fitted smooth function, in order to find the more likely functional form to fit the data (Hastie & Tibshirani 1990; Brenning 2008; Goetz et al. 2011). Specifically, it uses a link function to establish the relationship between the mean  $\mu$  of the response variables (probability of landslide occurrence) and a sum of a set of smooth functions of independent variables (Jia et al. 2008):

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

where  $g$  is the link function (logistic) and the  $f_i$  are smooth function (typically splines), each depends on a single explanatory variable  $x_i$  chosen in a set of  $n$  variables  $x_1 \dots x_n$ .

The GAM is based on the combination of linear and nonlinear smoothing functions and the application of different modelling policies according to the characteristics of the predictor factors. This allows to better describe the complex relationship between independent and response variables (Petschko et al. 2014).

The GAM is built using a stepwise variable selection. Starting from null model, each variable can be entered as linear (untransformed), nonlinear (non-parametrically transformed predictor of two equivalent degrees of freedom), or not included in the model. The minimization of Akaike information criterion (AIC) is used as selection criterion (Brenning 2008; Guisan et al. 2002; Goetz et al. 2011).

The implementation of the GAM was provided by means of the package ‘gam’ in R software (Hastie 2013).

The procedure adopted was subdivided in the following steps:

- (1) Multicollinear analysis to evaluate the dependence between the predictor variables and to comply with the basic assumptions of multiple regression models;
- (2) Selection of the most significant explanatory variables for the best estimation of the shallow landslide probability by using the AIC. This step is repeated within 100-fold bootstrap procedure of random selection of training and test data sets;
- (3) Evaluation of model forecasting capabilities and shallow landslide maps extraction.

The first step in the GAM implementation was the application of the multicollinear analysis. This step is a basic starting point, since the multicollinearity among explanatory variables is equivalent to have a redundancy of input information. Multicollinearity occurs when the predictor variables (in this case landslide predisposing factors) are linearly related with a strong correlation. The identification of this dependency is therefore necessary to reduce redundancy and to improve numerical stability in the subsequent analyses (Farrar & Glauber 1967). This analysis was performed with the function *Colldiag*, available in the R package *perturb*. This is an implementation of the regression collinearity diagnostic procedures found in Belsley et al. (1980). The procedure calculates the condition indexes of the matrix of the independent variables. A condition index higher than 30 (as suggest by Belsley et al. 1980) between a couple of variables indicates that those variables are not independent. Such predictors were excluded from the analyses in order to reduce collinearity.

In the second step, an equal number of non-landslide pixels was appended to the landslide database to avoid an over-estimation of non-landslide areas, according to a general method described by Dai and Lee (2002), Ayalew and Yamagishi (2005), Duman et al. (2006), and Mathew et al. (2009). Thus, the entire database, consisting of all landslide pixels and the same number of randomly selected non landslide pixels, was subdivided into 2 subsets: the training and the test sets. The training set, representing 2/3 of the dataset, was used to build the GAM fitting the samples, while the test set, including 1/3 of the dataset, was used to estimate the model accuracy. The process of training and test random selection was repeated in a 100-fold bootstrap procedure, allowing to identify the most frequent predictor variables. The variables whose frequency was more than 80% in the bootstrap extraction were identified as the most influent and used to build the model. Among the selected predictors, those linear and non-linear were identified. The discrimination was based on the higher percentage of selection obtained. The selected variables were used to build the model for the landslide occurrence prediction in all the study areas.

In the third step, the model forecasting capability was assessed through a repeated holdout method for regression with a binary response, slightly modified with respect to the standard procedure (Maindonald & Braun 2010). Holdout method is the simplest kind of cross-validation technique, allowing easy computational processes (McLachlan 1992; Molinaro et al. 2005). The repeated holdout is a k-fold repetition of holdout method, which consists of a random subsampling of different training and test sets, in proportion of 2/3 for testing and 1/3 for test. The repeated holdout method provides an estimate of the accuracy for 100 iterations calculated in all training and test sets. The results of accuracy were averaged to yield an overall accuracy and compared. Unlike the standard holdout method that separates the entire dataset into training and test sets, in this work these two sets were provided by the bootstrap model selection.

Another measure of the predictive performance was evaluated through the receiver operating characteristic (ROC) curve (Hosmer & Lemeshow 2000). In particular, the area under the ROC curve (AUROC) was compute to evaluate the model ability to discriminate landslide and non-landslide location. The ROC was calculated by plotting the sensitivity of a model (true positive rate) to the 1-specificity (false positive rate; Hosmer & Lemeshow 2000; Petschko et al. 2014). The AUROC can takes values from 0.5 (no discrimination) to 1.0 (perfect discrimination; Brenning 2005).

Moreover, the mean value of the 100 AUROC samples obtained from the 100-fold bootstrap procedure was calculated. Also the bootstrap 95% confidence bands of ROCs and the bootstrap 95% confidence intervals of AUROCs were obtained.

The 100 fitted bootstrap models were used to extend the prediction to the whole areas in order to obtain a distribution of landslide probability for each pixel. The mean values of each bootstrap distribution of 100 probability values were used to obtain the landslide susceptibility maps. A prediction uncertainty was associated to each estimated probability by computing the bootstrap 95% confidence intervals of landslide probability. The probability values were subdivided into 4 intervals in the susceptibility map:  $0 \leq p \leq 0.25$ ,  $0.25 < p \leq 0.50$ ,  $0.50 < p \leq 0.75$ ,  $0.75 < p \leq 1$ .

This classification method considers the equal probability interval of 0.25 for each susceptibility class. A value of 0.5 represents the same probability of occurrence or not of landslide. The ranges  $0 \leq p \leq 0.25$ ,  $0.25 < p \leq 0.50$ ,  $0.50 < p \leq 0.75$  and  $0.75 < p \leq 1$  indicate a low, medium-low, medium-high and high probability of landslide occurrence, respectively.

## 4. Results

### 4.1. Collinearity analysis and explanatory variables selection

In the Rio Frate, the Versa and the Pogliaschina catchments TRI and TWI are collinear explanatory variables and thus were excluded from the models. In the Vernazza catchment, in addition to these, CA was excluded as well.

**Table 3.** Absolute frequencies of explanatory variables (both linear and nonlinear) selected by 100-fold bootstrap procedure. NA: 'not available' indicates parameters discarded by the multicollinear analysis. In bold red were highlighted the explanatory variables selected by 100-fold bootstrap procedure with an absolute frequency greater or equal than 80. SL: Slope; TPI: Topographic position index; PLA: Planform curvature; PRO: Profile curvature; CS: Catchment slope; CA: Catchment area; RD: Distance from roads; ASP: Aspect; LU: Land use; GEO: Geology.

Catchments	SL	TPI	PLA	PRO	CS	CA	RD	ASP	LU	GEO
Rio Frate	<b>100</b>	37	59	78	59	12	12	<b>87</b>	63	13
Versa	<b>96</b>	41	62	64	<b>90</b>	45	18	<b>80</b>	<b>92</b>	72
Pogliaschina	<b>100</b>	<b>100</b>	<b>100</b>	72	<b>99</b>	49	9	<b>99</b>	<b>100</b>	<b>100</b>
Vernazza	17	<b>100</b>	31	69	63	NA	3	<b>98</b>	<b>100</b>	<b>100</b>

The frequency of predictor variables selected by the 100-fold bootstrap procedure was calculated. The ones with a frequency higher than 80% were selected (Table 3).

In the Rio Frate catchment only SL (100) and ASP (87) were considered. PRO and LU reached a high percentage (78 and 63, respectively), but fell under the selection threshold. In the Versa catchment, in addition to SL (96) and ASP (80), also CS (90) and LU (92) were insert in the model.

In the Vernazza catchment, the most influent predisposing factors were TPI, LU and GEO, with a frequency selection of 100, and ASP, with a frequency of 98. In the Pogliaschina catchment, the majority of the variables (GEO, LU, SL, ASP, PLA and TPI) reached a frequency of selection ranging between 99 and 100.

PRO, CA and RD, not reaching a frequency of selection greater than 80% in all the study areas, were excluded.

The selected continuous explanatory variables were distinguished into linear or nonlinear. The discrimination was based on the higher percentage of selection obtained in the bootstrap procedure (Table 4). In the Rio Frate and the Vernazza catchments the variables selected were linear. In the Versa and the Pogliaschina catchments all continuous explanatory variables were chosen as non-linear.

#### 4.2. Assessment of the predictive performance of the model

The models obtained in this work are characterized by a good predictive capability, as demonstrated by the mean accuracy and AUROC values of the 100 bootstrap iterations. For each catchment, the mean and standard deviation of accuracy were computed both in the training set and in test set. The results are quite similar, with an absolute differences ranging between 0.01 and 0.04 (Table 5).

A good predictive overall accuracy was also obtained by AUROC measures, with mean values ranging from 0.76 to 0.82 (Table 6). The best value was obtained for the Pogliaschina catchment. Besides, Table 6 highlights a small fluctuation of the bootstrap 95% confidence intervals of AUROCs

**Table 4.** Absolute frequencies of linear and non-linear selection for continuous explanatory variables. Each cell is split diagonally in two parts: in the right upper part and in left lower part are shown the frequencies of linear and non-linear choice of explanatory variables, respectively. The numbers highlighted in bold and italics are refer to the linear or non-linear type of explanatory variables that were chosen to build the model. NA: 'not available' indicates parameters discarded by model (selection frequency lower than 80).

Choice of linear and non-linear continuous explanatory variables				
<i>Explanatory</i>	<i>SL</i>	<i>TPI</i>	<i>PLA</i>	<i>CS</i>
<i>catchments</i>				
Rio Frate	<u>52</u> 48	NA	NA	NA
Versa	<u>66</u> 30	NA	NA	<u>43</u> 47
Pogliaschina	<u>81</u> 19	<u>68</u> 28	<u>72</u> 33	<u>50</u> 49
Vernazza	NA	<u>75</u> 25	NA	NA

**Table 5.** Mean and standard deviation of accuracy computed in the training and test sets.

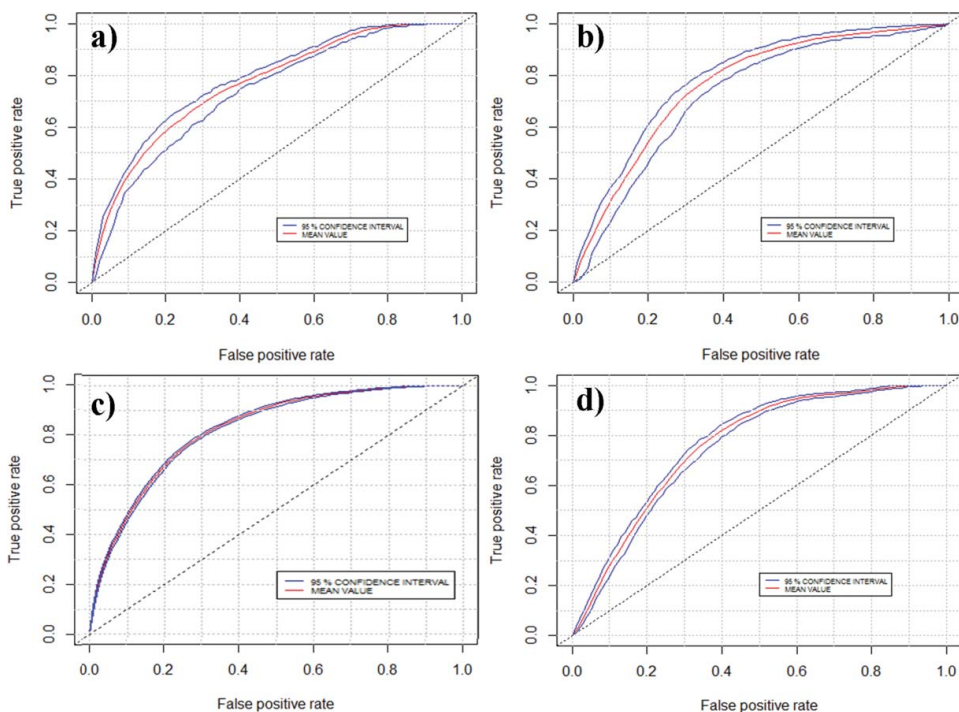
Catchment	Mean accuracy on training sets	Mean accuracy on test sets	Standard deviation of accuracy on training sets	Standard deviation of accuracy on test sets
Rio Frate	0.70	0.68	0.03	0.03
Versa	0.74	0.70	0.02	0.03
Pogliaschina	0.75	0.74	0.01	0.01
Vernazza	0.72	0.70	0.01	0.02

**Table 6.** Mean AUROC values obtained averaging output of 100 iterations for each study area and 95% bootstrap confidence intervals of AUROC.

Catchment	Mean AUROC	95 % confidence interval of AUROC
Rio Frate	0.77	0.75–0.78
Versa	0.76	0.74–0.77
Pogliaschina	0.82	0.82–0.85
Vernazza	0.76	0.74–0.77

with a maximum amplitude of 0.03. This result is also confirmed by the very narrow bootstrap 95% confidence bands of ROCs (Figure 6).

The maps showing the bootstrap 95% confidence intervals of landslide probability are illustrated in Figure 7. As it is possible to notice, the spatial variability of landslide probability is generally low in the great part of the catchments. The highest variability was observed in small portions of the Versa and the Pogliaschina catchments (Figure 7(b) and 7(c)), which was very concentrated in the Pogliaschina case. A larger fluctuation of middle values is however noticeable for the Rio Frate and the Vernazza catchments (Figure 7(a) and 7(d)). The overall results highlight that the landslide probability associated to a the majority of pixels values is reliable.

**Figure 6.** 95% bootstrap confidence bands of ROCs: (a) Rio Frate catchment; (b) Versa catchment; (c) Pogliaschina catchment; (d) Vernazza catchment.



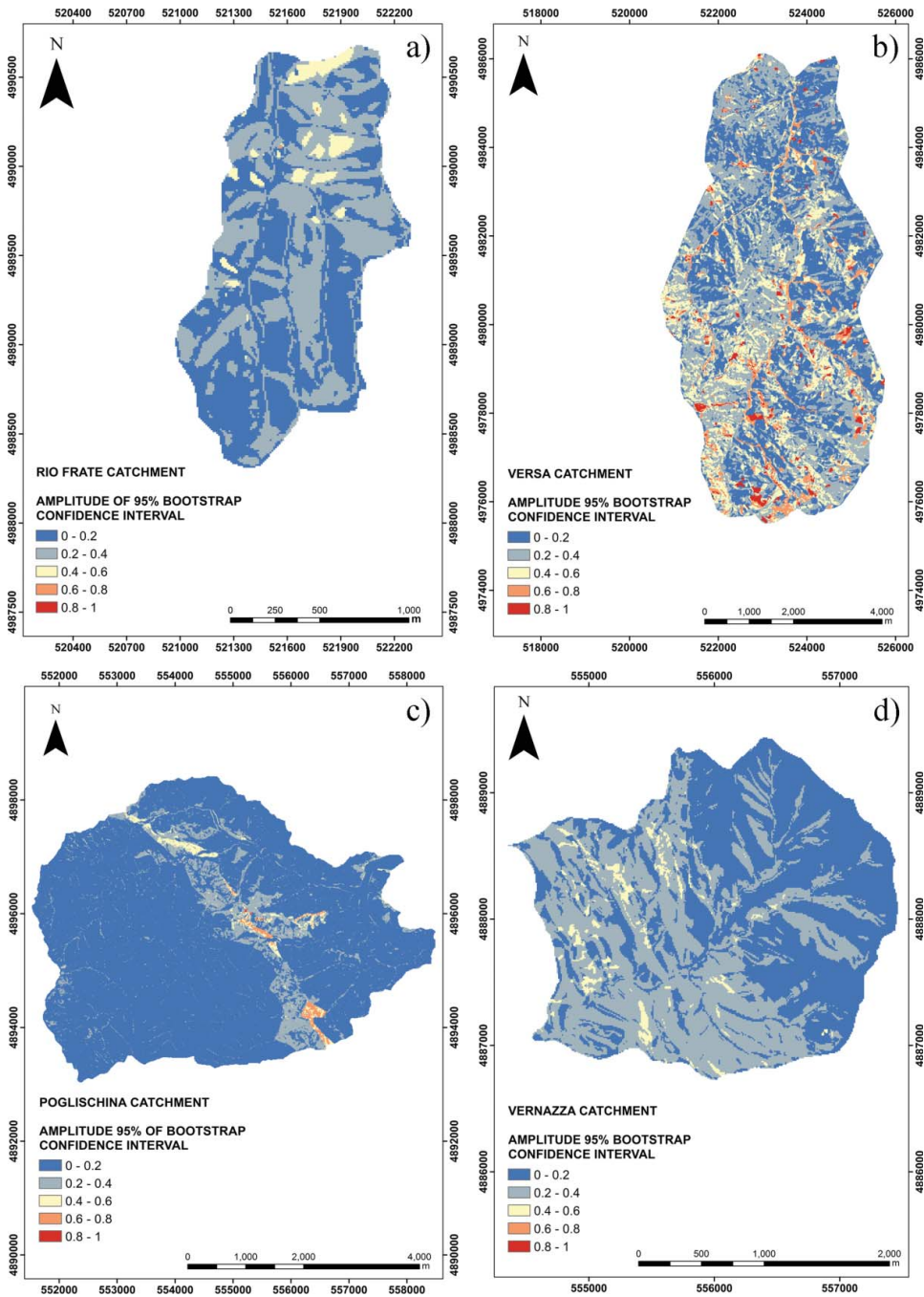
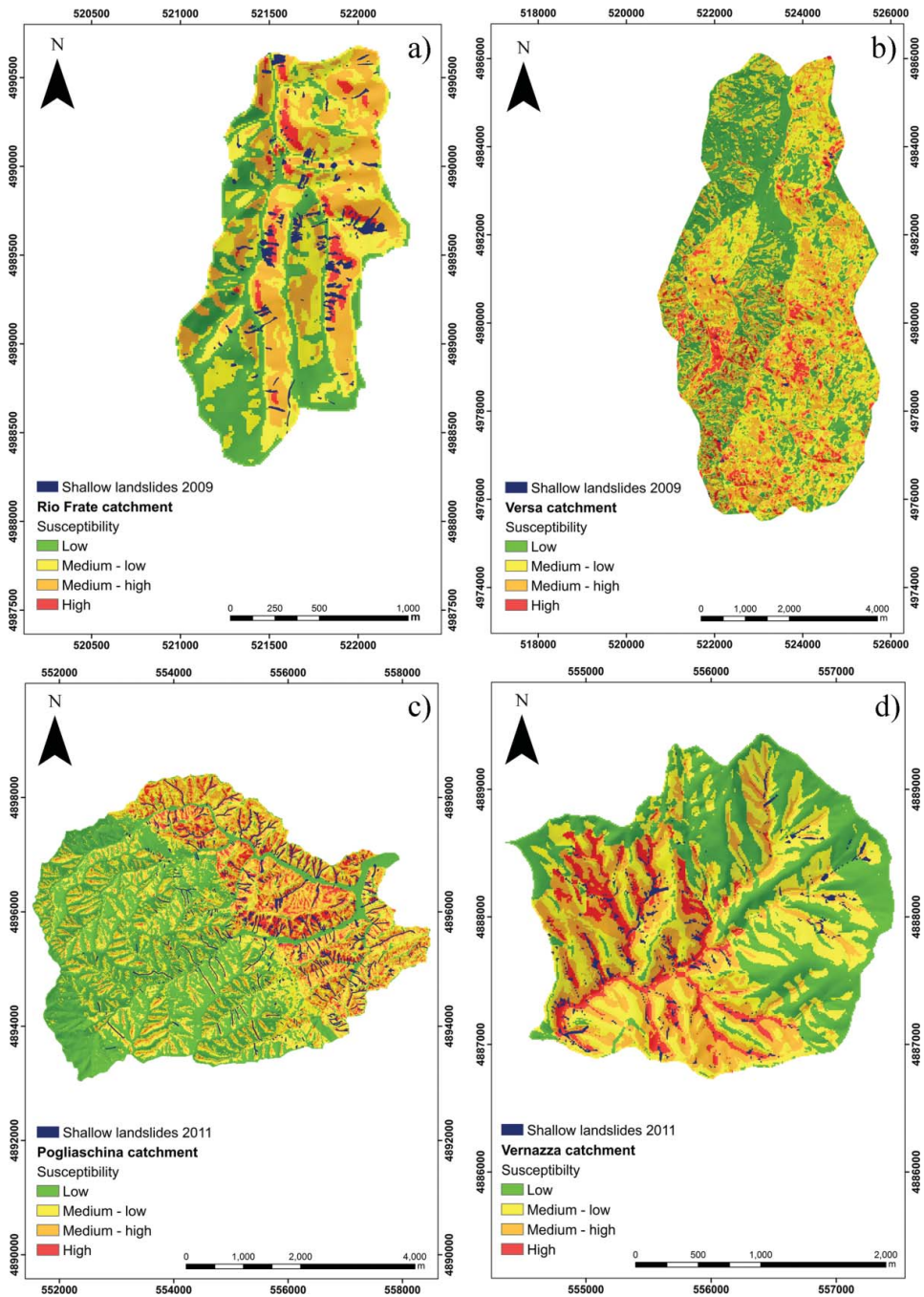


Figure 7. Maps of amplitude of 95% bootstrap confidence intervals of landslide probability: (a) Rio Frate catchment; (b) Versa catchment; (c) Pogliaschina catchment; (d) Vernazza catchment.

### 4.3. Shallow landslides susceptibility maps extraction

The shallow landslide susceptibility maps extracted by GAM are shown in Figure 8. In the Rio Frate catchment, 62% of the territory covers low and medium-low shallow landslide susceptibility classes, while the remaining 38% falls into medium-high and high shallow susceptibility classes (Figure 8(a)).



**Figure 8.** Shallow landslides susceptibility maps extracted after the GAM application: (a) Rio Frate catchment; (b) Versa catchment; (c) Pogliaschina catchment; (d) Vernazza catchment.

The highest shallow landslide susceptibility is localized in correspondence of steep slopes (range between 24 and 36°) and to East and South-West/West facing slopes.

The 66% of the area of the Versa catchment is classified with low and medium-low susceptibility, whereas the other 34% with medium-high and high susceptibility (Figure 8(b)). By comparing the shallow landslide susceptibility map with the more frequent variables, it is possible to observe that

the highest susceptibility is located in correspondence of areas with slopes comprised in the range  $18^{\circ}$ – $27^{\circ}$ . As in the Rio Frate catchment case, also in the Versa catchment, the slopes exposed towards South-West/West are more susceptible to landslide. In addition to the other geomorphological parameters, also the land use is very influent on slopes instability. In particular, cultivated areas (e.g. vineyards) cover the highest shallow landslide classes.

In the Pogliaschina catchment, a lower extension percentage (27.7%) of medium–high and high landslide susceptibility classes was obtained. Otherwise, 72.3% of territory fall into low and medium low landslide susceptibility classes (Figure 8(c)). It was found that concave slopes, South-East and South facing slopes and arenaceous Formations were more susceptible to landslide.

In the Vernazza catchment, 36% of the territory is located in medium–high and high shallow landslide susceptibility classes, whereas the remaining 64% has a low and medium–low probability of landslide (Figure 8(d)). In this case, the land use is the most influent parameter in landsliding. By comparing the areas with high values of susceptibility and the land-use map, it was possible to observe that slopes more prone to instability are situated in correspondence of terraced areas. In particular, the abandoned terraced areas are the most susceptible.

## 5. Discussion

This work was aimed to develop and test the capability of a comprehensive data-driven methodology, based on GAM, in order to assess the shallow landslide susceptibility of areas with different geological, geomorphological and environmental features.

In literature, the data-driven methods have been often applied in single areas (Guzzetti et al. 2006; Goetz et al. 2011; Galve et al. 2015), without investigating their predictive powerful in heterogeneous contexts. In this work, an innovative method for the evaluation of shallow landslide susceptibility of four different study areas was developed. This methodological strategy was able to identify the main geomorphological, hydrological, geological and land-use parameters controlling the shallow landslide occurrence, according to the characteristics of each investigated area.

The investigation of the most important predisposing factors could lead to improve the knowledge about mechanisms which regulate landslide occurrence. For this purpose, the collinearity analysis and 100-fold bootstrap technique were found to be objective tools to select the most influential variables.

Moreover, instead of using a single initiation point located in the upper part of landslides as response variable, the introduction in the model of a polygon representing the whole landslide source area (25% of the total landslide area) by means of an automatic procedure allowed to better characterize the detachment area. In addition, it allowed to reduce the spatial data quality degradation derived from the transformation from vector to raster. The degradation is in connection with the pixels size resolution, which could lead to a wrong characterization of shallow landslide initiation zone. At the size resolution of 10 meters, a single initiation point could almost entirely include portion of area outside the landslide perimeter during the process of conversion from vector to raster. Anyway, at this size resolution, also using 25% of the total landslide area revealed some problems. In case of landslide of small extent, an overestimation of landslide initiation areas occurred, wrongly including part of the accumulation areas.

As regards the model accuracy, the good forecasting capability of GAM was confirmed by high values of accuracy.

Nevertheless, the worst results were obtained in the Rio Frate catchment, where the GAM wasn't completely able to identify roto-translational slides developed in correspondence of road cut scarps. This is mainly due to the 10 meters spatial resolution of input data, which was not capable to take into account all detailed terrain information. Moreover, the RD variable was excluded during the processing phase. This lack of information could lead to a misleading analysis of those areas where local peculiar factors may affect slope instability (Figure 9).



 Shallow landslides not identified

### Rio Frate catchment

#### Susceptibility

 Medium - high

 High

**Figure 9.** Example of areas within Rio Frate catchment where the model has not correctly predicted the presence of roto-translational slides developed in the road cut scarps. The aerial photograph of the area was taken by Ditta Rossi s.r.l. (Brescia, Italy) on 18 May 2009.

In the Rio Frate catchment case, the slope is the variable that mostly contributes to instability. The analysis of the landslide-slope class distribution showed that the majority of shallow landslides occurred on steep slopes, which can reach values higher than  $35^\circ$ . Besides, the most susceptible slopes were exposed toward South-West/West. As observed by other authors (van Westen et al. 2008; Demir et al. 2013), the slope direction reflects differences in rainfall amount, solar radiation and soil moisture conditions, which could favour differently the landslide triggering.

In the Versa catchment, in addition to the geomorphological parameters, also the cultural practices were influent in shallow landslides occurrence. In particular, the vineyards were the land-use

class most affected by the highest probability to landslide. This fact was confirmed by the landslide inventory related to the 27–28 April 2009, where the majority of shallow landslides were located within cultivated areas (55% within vineyards).

In the Vernazza catchment, the land use played an important role on shallow landslide susceptibility. As already observed by Cevasco et al. (2014) and Galve et al. (2015), the shallow landslide susceptibility tends to increase in abandoned terraces (with poor vegetation cover) and thus also in terraces which have been abandoned for a long time (with dense vegetation cover). The lowest values of susceptibility were found in wood and scrub land. This fact highlights that the increase of susceptibility is predominant on areas characterized by land abandonment. Effectively, in the Vernazza catchment, from 1960 to 2011, strongly modification of agricultural practices occurred. A high percentage of terraced areas, especially cultivated with vineyards and olive groves were abandoned. On 25 October 2011, most of the landslides occurred in correspondence of abandoned terraces. This is probably in relation to the lack of maintenance of the dry-stone walls, which might have favoured instability phenomena (Cevasco et al. 2013a; Galve et al. 2015).

In the Pogliaschina catchment, despite the land use resulted important, the influence of geomorphological parameters and the geology on the GAM predictive performance is high. Geology and planform curvature are the most relevant variables. In fact, the highest susceptibility was identified in correspondence of concave underlain by arenaceous Formations.

## 6. Conclusions

The proposed methodology based on the GAM technique led to assess the shallow landslide susceptibility in different environmental contexts. In particular, it was able to better characterize the shallow landslide predisposition of four territories with different geological, geomorphological, land-use characteristics and shallow landslide types.

In this work, the use of landslide inventories referred only to a single event has represented the major limitation. In absence of multi-temporal landslide database, we randomly partitioned the entire dataset in two parts (training and test subsets) within a 100-fold bootstrap procedure. The model accuracy was carried out for each bootstrap procedure, comparing results of training subset and test subset (landslide pixels not used to fit the model).

Despite this limit, the GAM performance was characterized by a good forecasting capability, as demonstrated by the accuracy and AUROC values of the 100 bootstrap iterations, which values ranging from 0.70 to 0.75 and from 0.76 to 0.82, respectively.

In conclusion, the novelty of this work is to present a processing chain characterized by simplicity, reproducibility and predictive efficiency in the assessment of landslide susceptibility of different territories, despite the influence of different rainfall conditions and environmental settings. The selection of the most significant variables provided a better description of the shallow landslide occurrence and distribution.

Moreover, the use of input data simply derived from DEMs allows to obtain a good level of accuracy and predictive efficiency also in case of lack of exhaustive information. This could permit the use of this methodology also in contexts where is difficult to obtain effective monitoring tools.

In the future development, the presented methodology will be applied also in monsoon-influenced areas where, despite the high vulnerability that characterizes those areas, there is still lack of comprehensive study concerning the landslide susceptibility assessment. This will allow to deeply assess the flexibility and repeatability of the developed processing chain, comparing the results with the ones obtained from the study areas analysed in this research and characterized by climatic, geological and geomorphological conditions completely different from the ones typical of tropical areas.


## Acknowledgment

The authors thank the anonymous reviewers for their suggestions and contribution to the work.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## ORCID

Maria Giuseppina Persichillo  <http://orcid.org/0000-0002-3041-9393>

Claudia Meisina  <http://orcid.org/0000-0003-3673-3794>

## References

- Abbate E, Fanucci F, Benvenuti M, Bruni P, Cipriani N, Falorni P, Fazzuoli M, Morelli D, Pandeli E, Papini M, et al. 2005. Carta Geologica d'Italia alla scala 1:50.000, Foglio 248, La Spezia. [Geological Map of Italy at scale 1:50.000. Sheet 248, La Spezia]. Serv. Geol. d'Italia. Roma. Available from: [http://www.isprambiente.gov.it/Media/carg/248\\_LASPEZIA/Foglio.htm](http://www.isprambiente.gov.it/Media/carg/248_LASPEZIA/Foglio.htm)
- Agenzia Regionale per la Protezione dell'Ambiente Ligure-Centro Funzionale Meteorologico di Protezione Civile della Regione Liguria. 2012. Uno tsunami venuto dai monti - Provincia della Spezia 25 ottobre 2011 rapporto di evento meteo-idrologico [Tsunami came from the mountains - La Spezia Province October 25, 2011 report of hydro-meteorological event]. Quaderni ARPAL. 1:1–96.
- Ayalew L, Yamagishi H. 2005. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. *Geomorphology*. 65:12–31.
- Baeza C, Corominas J. 2001. Assessment of shallow landslide susceptibility by means of multivariate statistical techniques. *Earth Surf Processes Landforms*. 26:1251–1263.
- Baeza C, Lantada N, Moya J. 2010. Validation and evaluation of two multivariate statistical models for predictive shallow landslide susceptibility mapping of the Eastern Pyrenees (Spain). *Environ Earth Sci*. 61:507–523.
- Bartelletti C, D'Amato Avanzi G, Galanti Y, Giannecchini R, Mazzali A. 2015. Assessing shallow landslide susceptibility by using the SHALSTAB model in Eastern Liguria (Italy). *Rend Online Soc Geol It*. 35:17–20.
- Bortolotti V, Principi G, Abbate E, Chiari M, Aiello IW, Mannori G, Reale V, Pandeli E. 2011. Carta Geologica d'Italia alla scala 1:50.000, Foglio 232, Sestri Levante [Geological Map of Italy at scale 1:50.000, Sheet 232, Sestri Levante]. Serv. Geol. d'Italia. Roma. Available from: [http://www.isprambiente.gov.it/Media/carg/232\\_SESTRI\\_LEVANTE/Foglio.html](http://www.isprambiente.gov.it/Media/carg/232_SESTRI_LEVANTE/Foglio.html)
- Belsley DA, Kuh E, Welsch RE. 1980. *Regression diagnostics: identifying influential data and sources of collinearity*. New York: John Wiley and Sons.
- Brabb EE. 1984. Innovative approaches to landslide hazard mapping. In: 4th Int. Symp. on Landslides, Toronto, Canada. 1:307–324.
- Brandolini P, Faccini F, Robbiano A, Terranova R. 2007. Geomorphological hazards and monitoring activity along the western rocky coast of the Portofino Promontory (Italy). *Quat Int*. 171–172:131–142.
- Brandolini P. 2016. The outstanding terraced landscape of the Cinque Terre coastal slopes (Eastern Liguria). In: Soldati M, Marchetti M, editors. *Landforms and landscapes of Italy*. Switzerland: Springer International Publishing.
- Brenning A. 2005. Spatial prediction models for landslide hazards: review, comparison and evaluation. *Nat Hazards Earth Syst Sci*. 5:853–862.
- Brenning A. 2008. Statistical geocomputing combining R and SAGA: the example of landslide susceptibility analysis with generalized additive models. In: Böhner J, Blaschke T, Montanarella L, editors. *SAGA — Seconds Out (= Hamburger Beiträge zur Physischen Geographie und Landschaftsökologie, 19)*. Hamburg; p. 23–32.
- Brenning A, Schwinn M, Ruiz-Páez AP, Muenchow J. 2015. Landslide susceptibility near highways is increased by one order of magnitude in the Andes of southern Ecuador, Loja province. *Nat Hazards Earth Syst Sci*. 15:45–57.
- Caine N. 1980. The rainfall intensity-duration control of shallow landslides and debris flows. *Geogr Ann*. 62:23–27.
- Campus S, Forlati F, Susella G, Tamberlani F. 1998. Frane per mobilizzazione delle coperture detritiche [Landslides for mobilization of debris covers]. In: Regione Piemonte, editor. *Eventi alluvionali in Piemonte*. Turin; p. 266–287.
- Carobene L, Cevasco A. 2011. A large scale lateral spreading, its genesis and Quaternary evolution in the coastal sector between Cogoleto and Varazze (Liguria – Italy). *Geomorphology*. 129:398–411.
- Catani F, Casagli N, Ermini L, Righini G, Menduni G. 2005. Landslide hazard and risk mapping at catchment scale in the Arno River Basin. *Landslides*. 2:329–342.
- Catani F, Lagomarsino D, Segoni S, Tofani V. 2013. Landslide susceptibility estimation by random forests technique: sensitivity and scaling issues. *Nat Hazards Earth Syst Sci*. 13:2815–2831.
- Cevasco A, Brandolini P, Scopesi C, Rellini I. 2013a. Relationships between geo-hydrological processes induced by heavy rainfall and land-use: the case of 25 October 2011 in the Vernazza catchment (Cinque Terre, NW Italy). *J Maps*. 9:289–298.

- Cevasco A, Pepe G, Brandolini P. 2013b. Geotechnical and stratigraphic aspects of shallow landslides at Cinque Terre (Liguria, Italy). *Rend Online Soc Geol It.* 24:52–54.
- Cevasco A, Pepe G, Brandolini P. 2014. 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.
- Cevasco A, Diodato N, Revellino P, Fiorillo F, Grelle G, Guadagno FM. 2015. Storminess and geo-hydrological events affecting small coastal basins in a terraced Mediterranean environment. *Sci Total Environ.* 532:208–219.
- Chen Z, Wang J. 2007. Landslide hazard mapping using logistic regression model in Mackenzie Valley, Canada. *Nat Hazards.* 42:75–89.
- Corominas J, van Westen C, Frattini P, Cascini L, Malet J-P, Fotopoulou S, Catani F, Van Den Eeckhaut M, Mavrouli O, Agliardi F, et al. 2014. Recommendations for the quantitative analysis of landslide risk. *Bull Eng Geol Environ.* 73:209–263.
- Corradi N, Cevasco A, Iannotta MA, Cavallo C, Schiaffino C, Brignone M, Ferrari M. 2013. Le coste della Liguria [The Liguria's coasts]. In: Abis A, Campo C, Bonaria Careddu M, Deriu M, editors. *La rete per il monitoraggio delle spiagge. Applicazioni ed esperienze sviluppate nel progetto ResMar (Azione di sistema A)*. Olbia: Taphros Ed.; p. 41–59.
- Costanzo D, Rotigliano E, Irigaray C, Jiménez-Perálvarez JD, Chacón J. 2012. Factors selection in landslide susceptibility modelling on large scale following the gis matrix method: application to the river Beiro basin (Spain). *Nat Hazards Earth Syst Sci.* 12:327–340.
- Crosta GB, Agliardi F, Frattini P, Sosci R. 2012. SafeLand Deliverable 1.1: landslide triggering mechanisms in Europe – overview and state of the art. Identification of mechanisms and triggers. Available from: <http://www.safeland-fp7.eu/>.
- Cruden DM, Varnes DJ. 1996. Landslide types and processes. In: Turner AK, Schuster RL, editors. *Landslides: investigation and mitigation*. Sp. Rep. 247. Washington (DC): Transportation Research Board, National Research Council. National Academy Press; p. 36–75.
- D'Amato Avanzi G, Galanti Y, Giannecchini R, Bartelletti C. 2014. Shallow landslides triggered by the 25 October 2011 extreme rainfall in Eastern Liguria (Italy). In: Lollino G, Giordan D, Crosta GB, Corominas J, Azzam R, Wasowski J, Sciarra N, editors. *Engineering geology for society and territory*. Switzerland: Springer International Publishing; p. 515–519.
- Dai FC, Lee CF, Ngai YY. 2002. Landslide risk assessment and management: an overview. *Eng Geol.* 64:65–87.
- Dai FC, Lee CF. 2002. Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. *Geomorphology.* 42:213–228.
- De Vita P, Cevasco A, Cavallo C. 2012. Detailed rock failure susceptibility mapping in steep rocky coasts by means of non-contact geospatial surveys: the case study of the Tigullio Gulf (Eastern Liguria, Northern Italy). *Nat Hazards Earth Syst Sci.* 12:867–880.
- Demir G, Aytikin M, Akgün A, İkizler SB, Tatar O. 2013. A comparison of landslide susceptibility mapping of the eastern part of the North Anatolian Fault Zone (Turkey) by likelihood-frequency ratio and analytic hierarchy process methods. *Nat Hazards.* 65:1481–1506.
- Devkota KC, Regmi AD, Pourghasemi HR, Yoshida K, Pradhan B, Ryu I, Dhital MR, Althuwaynee O. 2013. Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling-Narayanghat road section in Nepal Himalaya. *Nat Hazards.* 65:135–165.
- Duman TY, Can T, Gokceoglu C, Nefeslioglu HA, Sonmez H. 2006. Application of logistic regression for landslide susceptibility zoning of Cekmece Area, Istanbul, Turkey. *Environ Geol.* 51:241–256.
- Farrar DE, Glauber RR. 1967. Multicollinearity in regression analysis: the problem revisited. *Rev Econ Stat.* 49:92–107.
- Fell R, Corominas J, Bonnard C, Cascini L, Leroi E, Savage WZ (on behalf of the JTC-1 Joint Technical Committee on Landslides and Engineered Slopes). 2008. Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. *Eng Geol.* 102:85–98.
- Galve JP, Cevasco A, Brandolini P, Soldati M. 2015. Assessment of shallow landslide risk mitigation measures based on land use planning through probabilistic modelling. *Landslides.* 12:101–114.
- Giannecchini R. 2006. Relationship between rainfall and shallow landslides in the southern Apuan Alps (Italy). *Nat Hazards Earth Syst Sci.* 6:357–364.
- Giannecchini R, Galanti Y, D'Amato Avanzi G. 2012. Critical rainfall thresholds for triggering shallow landslides in the Serchio River Valley (Tuscany, Italy). *Nat Hazards Earth Syst Sci.* 12:828–842.
- Giannecchini R, Galanti Y, D'Amato Avanzi G, Barsanti M. 2016. Probabilistic rainfall thresholds for triggering debris flows in a human-modified landscape. *Geomorphology.* 257:94–107.
- Goetz JN, Guthrie RH, Brenning A. 2011. Integrating physical and empirical landslide susceptibility models using generalized additive models. *Geomorphology.* 129:376–386.
- Goetz JN, Brenning A, Petschko H, Leopold P. 2015. Evaluating machine learning and statistical prediction techniques for landslide susceptibility modelling. *Comput Geosci.* 81:1–11.
- Guisan A, Edwards TC, Hastie T. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecol Modell.* 157:89–100.

- Guzzetti F, Carrara A, Cardinali M, Reichenbach P. 1999. Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. *Geomorphology*. 31:181–216.
- Guzzetti F, Galli M, Reichenbach P, Ardizzone F, Cardinali M. 2006. Landslide hazard assessment in the Collazzone area, Umbria, Central Italy. *Nat Hazards Earth Syst Sci*. 6:115–131.
- Hastie TJ, Tibshirani R. 1990. *Generalized additive models*. London: Chapman & Hall.
- Hastie T. 2013. *GAM: Generalized Additive Models*. R package version 1.09. Available from: <http://CRAN.R-project.org/package=gam>.
- Hosmer DW, Lemeshow S. 2000. *Applied logistic regression*. New York: Wiley.
- Hungr O, Evans SG, Bovis M, Hutchinson JN. 2001. Review of the classification of landslides of the flow type. *Environ Eng Geosci*. 7:221–238.
- Hungr O, Leroueil S, Picarelli L. 2014. The Varnes classification of landslide types, an update. *Landslides*. 1:167–194.
- Jia G, Yuan T, Liu Y, Zhang Y. 2008. 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.
- Maindonald J, Braun WJ. 2010. *Data analysis and graphics using R: an example-based approach*. Cambridge Series in Statistical and Probabilistic Mathematics; p. 552.
- Mathew J, Jha VK, Rawat GS. 2009. Landslide susceptibility zonation mapping and its validation in part of Garhwal Lesser Himalaya, India, using binary logistic regression analysis and receiver operating characteristic curve method. *Landslides*. 6:17–26.
- McLachlan GJ. 1992. *Discriminant analysis and statistical pattern recognition*. New York: John Wiley & Sons.
- Molinario AM, Simon R, Pfeiffer RM. 2005. Prediction error estimation: a comparison of resampling methods. *Bioinformatics*. 21:3301–3307.
- Montgomery DR, Dietrich WE. 1994. A physically based model for the topographic control on shallow landsliding. *Water Resour Res*. 30:1153–1171.
- Montrasio L, Valentino R. 2007. Experimental analysis and modelling of shallow landslides. *Landslides*. 4:291–296.
- Pereira S, Zêzere JL, Bateira C. 2012. Technical Note: Assessing predictive capacity and conditional independence of landslide predisposing factors for shallow landslide susceptibility models. *Nat Hazards Earth Syst Sci*. 12:979–988.
- Petschko H, Brenning A, Bell R, Goetz J, Glade T. 2014. Assessing the quality of landslide susceptibility maps – case study Lower Austria. *Nat Hazards Earth Syst Sci*. 14:95–118.
- Phillips JD. 2003. Sources of nonlinearity and complexity in geomorphic systems. *Prog Phys Geog*. 27:1–23.
- Phillips JD. 2006. Evolutionary geomorphology: thresholds and nonlinearity in landform response to environmental change. *Hydrol Earth Syst Sci*. 10:731–742.
- Puccinelli A, D'Amato Avanzi G, Perilli N. 2015. Carta Geologica d'Italia alla scala 1:50.000, Foglio 233, Pontremoli. [Geological Map of Italy at scale 1:50.000, Sheet 233, Pontremoli]. *Serv. Geol. d'Italia*. Roma. Available from: [http://www.isprambiente.gov.it/Media/carg/233\\_PONTREMOLI/Foglio.html](http://www.isprambiente.gov.it/Media/carg/233_PONTREMOLI/Foglio.html).
- Quinn P, Beven K, Chevallier P, Planchon O. 1991. The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. *Hydrol Processes*. 5:59–79.
- Raggi G. 1985. Neotettonica ed evoluzione paleogeografica PlioPleistocenica del bacino del fiume Magra [Neotectonics and Plio Pleistocene paleogeographic evolution of the Magra river basin], *Memorie della Società Geologica Italiana*. 30:35–62.
- Regione Liguria. 2006. Carta geologica, La Spezia. Tavoleta n° 248.4, scala 1:25.000 [Geological map, La Spezia. 1 sheet :1:25.000].
- Regione Liguria. 2013. Carta di uso del suolo. Scala 1:10.000. Edizione 2012. [Land use map. 1 sheet :1:10.000. 2012 edition].
- Riley SJ, DeGloria SD, Elliot R. 1999. A terrain ruggedness index that quantifies topographic heterogeneity. *Intermountain J Sci*. 5:3–27.
- Sidle RC, Ochiai H. 2006. *Landslides: processes, prediction, and land use*. Water Resources Monograph 18. Washington (DC): American Geophysical Union.
- Seibert J, Stendahl J, Sørensen R. 2007. Topographical influences on soil properties in boreal forests. *Geoderma*. 141:139–148.
- Soeters R, Van Westen CJ. 1996. Slope instability recognition, analysis and zonation. In: Turner AK, Schuster RL, editors. *Landslides investigation and mitigation*. TRB Special Report 247. Washington (DC): National Academy Press; p. 129–177.
- Tasser E, Mader M, Tappeiner U. 2003. Effects of land use in alpine grasslands on the probability of landslides *Basic Appl Ecol*. 4:271–280.
- Terranova R, Zanzucchi G, Bernini M, Brandolini P, Campobasso S, Faccini F, Zanzucchi F. 2006. Geologia, geomorfologia e vini del Parco Nazionale delle Cinque Terre (Liguria, Italia) [Geology, geomorphology and wines of the National Park of the Cinque Terre (Liguria, Italy)]. *Bollettino della Società Geologica Italiana. Spec.* 6:115–128.
- van Westen CJ, Castellanos E, Kuriakose SL. 2008. Spatial data for landslide susceptibility, hazard, and vulnerability assessment: an overview. *Eng Geol*. 102:112–131.
- Varnes DJ. 1984. *Landslide hazard zonation – a review of principles and practice*. Paris: UNESCO; p. 63.
- Vercesi P, Scagni G. 1984. Osservazioni sui depositi conglomeratici dello sperone collinare di Stradella. *Rendiconti Società Geologica Italiana*. 7:23–26.



- Weiss AD. 2001. Topographic position and landforms analysis [dissertation]. ESRI User Conference, Poster presentation. San Diego (CA).
- Zevenbergen LW, Thorne CR. 1987. Quantitative analysis of land surface topography. *Earth Surf Processes Landforms*. 12:47–56.
- Zizioli D, Meisina C, Valentino R, Montrasio L. 2013. Comparison between different approaches to modelling shallow landslide susceptibility: a case history in Oltrepo Pavese, Northern Italy. *Nat Hazards Earth Syst Sci*. 13:559–573.