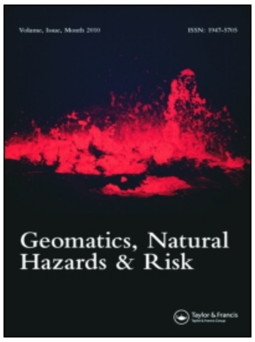

Taylor & Francis
Taylor & Francis Group

Geomatics, Natural Hazards and Risk

ISSN: 1947-5705 (Print) 1947-5713 (Online) Journal homepage: http://www.tandfonline.com/loi/tgnh20

The artificial neural network for the rockfall susceptibility assessment. A case study in Basilicata (Southern Italy)

Lucia Losasso & Francesco Sdao

To cite this article: Lucia Losasso & Francesco Sdao (2018) The artificial neural network for the rockfall susceptibility assessment. A case study in Basilicata (Southern Italy), Geomatics, Natural Hazards and Risk, 9:1, 737-759, DOI: 10.1080/19475705.2018.1476413

To link to this article: https://doi.org/10.1080/19475705.2018.1476413

Published online: 25 Aug 2018.

GEOMATICS, NATURAL HAZARDS AND RISK
2018, VOL. 9, NO. 1, 737–759
https://doi.org/10.1080/19475705.2018.1476413

The artificial neural network for the rockfall susceptibility assessment. A case study in Basilicata (Southern Italy)

Lucia Losasso and Francesco Sdao

School of Engineering, University of Basilicata, Potenza, Italy

ABSTRACT

This paper presents the results obtained by the elaboration of an artificial neuronal network for the creation of a rockfall susceptibility map. The analysis was carried out by analysing the predisposing and triggering factors of the rockfall phenomenon. The parameters considered for this study and representing the input data of the artificial neural network are factors such as: gradient, soil use, lithology, rockfall source areas and kinetic energy values obtained by considering the probable pathways of the blocks through simulations with dedicated softwares, DEMs and niches of the rockfalls that have already occurred in the past. The processing of this data (required in a versatile dedicated software for the realization of the artificial neural network in ASCII format) is done using GIS softwares, useful tools for the creation of hazard maps. An important step is the realization of the rockfall inventory map: it allows to identify the training set (consisting of 50% of the pixels relative to the rockfall niches) for the network training and the testing set (considering the remaining 50% of the pixels relative to the rockfall niches) to assess the network accuracy by overlaying the rockfall niches belonging to the testing set with the obtained susceptibility map.

ARTICLE HISTORY
Received 27 December 2017
Accepted 2 May 2018

KEYWORDS
Analysis Neural Networks (ANNs); Basilicata; GIS; Italy; rockfall susceptibility map

1. Introduction

The rockfalls represent a particularly widespread problem in Italy. The gravity of these phenomena is tied not only to the risks for people's safety, but also to the enormous damage they can cause to inhabited centres and to linear communication infrastructures.

The assessment of the mass movement hazard and risk, with particular reference to rockfalls, is a topic well-felt by the international scientific community and by the territorial management authorities because of the speed and the unpredictability with which they can occur. In particular, rockfalls develop mainly in mountainous and hilly areas which often cause serious damages to the population and to the exposed elements. The warning signs that announce an imminent detachment or movement are often, at least initially, slow and subtle except through instrumental observation, so only an assiduous control of

CONTACT Lucia Losasso lucia.losasso@unibas.it

the territory allows to obtain such precious information that, even though they can't give an accurate date of the event, provide clues about the possible accelerations of movements and deformations in progress. As it is well known, it is not possible, at the present state of knowledge, to simultaneously provide detailed information about the time and place the phenomenon can occur. However, it is possible to know a priori where the most dangerous areas are located, thanks to the analysis of the many predisposing factors and the triggering ones. Assessing the landslide risk of an area is a necessary and indispensable operation for proper land management and conservation. This step represents the first phase of the study for the purpose of assessing and drafting thematic mapping of landslide risk. The choice of the scale and of the detail degree are the first elements to consider when starting a geological and geomorphological study of landslide susceptibility (Losasso, Jaboyedoff et al. 2017). In particular, this choice is tied both to the type of the available data useful for the assessment of susceptibility, and to the methodology of the analysis to be used.

The regional scale makes possible to focus on more local problems than the national scale, although its use is still reserved for studies involving large territories, extending for thousands of kilometre squares. These studies are usually carried out using aggregate data to identify large geomorphological units presenting some structural, geological and morphological uniformity. Regional-level studies are also useful primarily for statistical purposes, in the process of defining general plans for emergency control and in the process of planning the resources at national level.

The hazard is, as is well known, the characterization of the unpredictability of a phenomenon with certain characteristics. This evaluation is generally complex and requires quantification, both at spatial and temporal levels, of the probability of occurrence of the phenomenon. More specifically, the landslide susceptibility provides the assessment of the probability that a phenomenon occurs in a given area.

Spatial hazard maps can be realized using different modelling approaches (Carrara et al. 1977, 1991; Carrara 1983; van Westen 1993; van Westen and Terlien 1996; Guzzetti et al. 1999; Ercanoglu and Temiz 2011; Lee et al. 2012) with different complexity degrees (Figures 1 and 2).

In particular, it is possible to adopt qualitative methods, using direct heuristic approaches based on a careful analysis of the territory (through field surveys, archive data collection, etc.), in order to analyse existing or previous phenomena, or indirect heuristic approaches based on the knowledge of the mass movement distribution in the area under consideration and of the predisposing or triggering factors (to every factor it is assigned a weight according to its importance within the process).

The quantitative approaches, however, can be led back to statistical models (Ayalew and Yamagishi 2005), probabilistic models (Chung and Fabbri 1999) and may also refer to Soft Computing methods such as Artificial Neuronal Networks. The assessment of the rockfall spatial susceptibility is a particularly important process since it is essential to be able to assess in advance the predisposition of a slope to collapse.

The main objective of the research in this paper is to provide an example of how the proposed approach could be used. Specifically, a map of the rockfall spatial hazard in a sample area of the Province of Potenza territory, heavily affected by rockfalls, especially along the communication lines has been created. In fact, the SP13

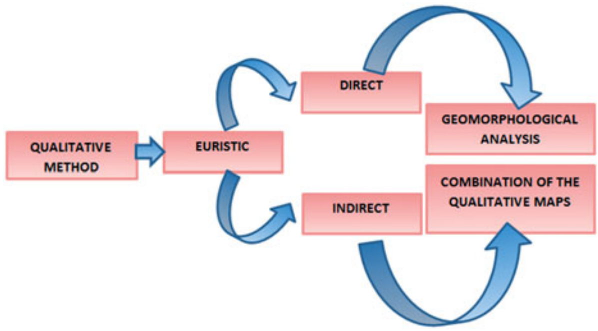

Figure 1. Qualitative methods to assess the mass movement hazard.

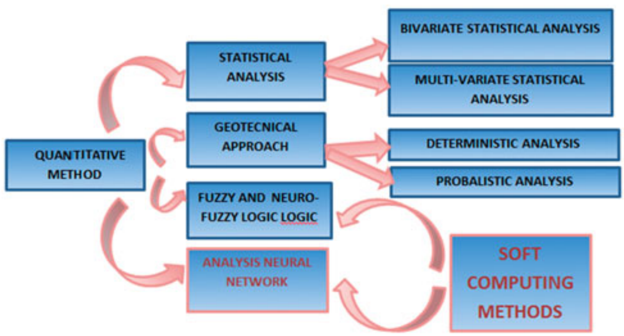

Figure 2. Quantitative methods to assess the mass movement hazard.

connecting the SS Basentana with two villages of growing touristic importance, Castelmezzano and Pietrapertosa, has been closed several times because of the collapse of large rocky blocks from the adjacent slopes.

2. Study area

2.1. Geological characteristics and types of involved materials in a zone of the Province of Potenza territory

The study area lies with the side of the Lucan Apennine in front of the valley of the Basento river.

The importance of assessing the susceptibility level due to natural phenomena in this area is related to the presence of numerous mass movements and landslides many of which are due to rockfalls (Figure 3(a)), often causing the closure of the roads and the partial isolation of the small villages that arise right in the heart of the Dolomites: Castelmezzano and Pietrapertosa (Figure 3(c)) (Losasso, Pascale et al. 2017).

The Apennine Chain constitutes a fold orogene originated by the Upper Oligocene–Myocene (Figure 3(b)). The Basilicata Region occupies the central part of the southern section of the Apennine Chain, which is known in the specialist literature as Lucan Apennines (Figure 4).

The area in the analysed territory (province of Potenza) from the geological point of view falls in the north-western sector of Sheet n. 490 'Stigliano', scale 1:50,000 of the

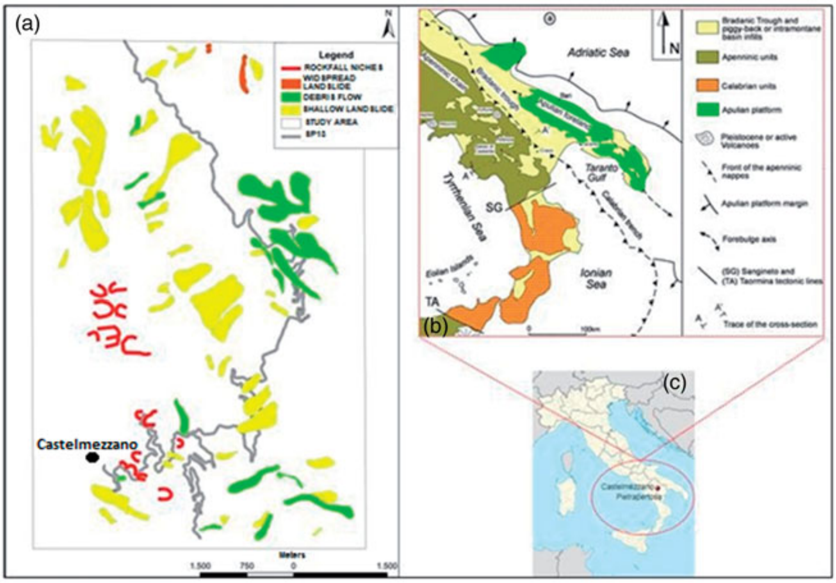

Figure 3. (a) Study area of about 42 km^2 falling in the territory of Castelmezzano and Pietrapertosa villages: landslide inventory map (Sdao and Sole, 2013); (b) Geological map of Southern Italy (Bentivenga et al., 2015); (c) Location of the study area.

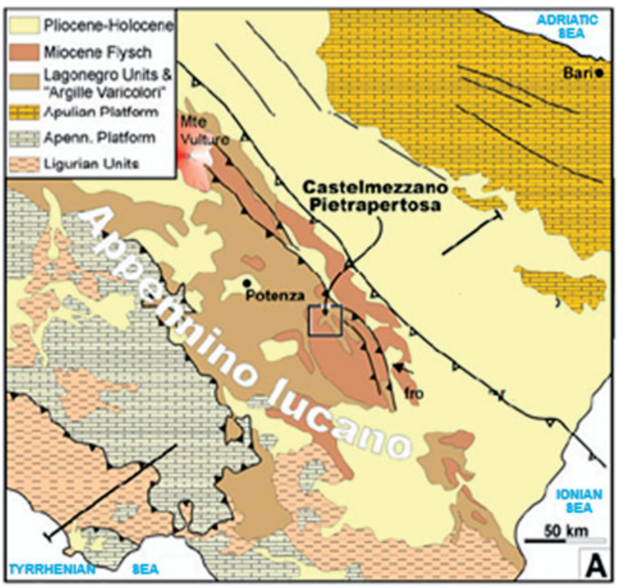

Figure 4. Detailed geological-structural scheme of the Lucan Apennines.

Geological Map of Italy. This area is made up of arenaceous-clay rocks linked to the Gorgoglione Flysch, represented by a dense alternation of turbiditic sandstone with intertwining of siltite and grey-blue clays. The arenaceous benches are well stratified and very slatted. In some places, especially near Pietrapertosa, there is a dense alternation of clay and silty clay complex with arenaceous intercalation. The Gorgoglione Flysh is disposed in a very inclined monoclinal characterized by NO-SE orientation and inclination of 40°/50°. Along the analysed slope, the layers show an orientation at an oblique angle to slope and hogback morphological shapes. The slope steepness, coupled with the intense stratification and cracking of the arenaceous benches, makes the area particularly fragile from the morphological point of view which is revealed by collapsing phenomena. From a morphological point of view, the territory under consideration has typical morphological characteristics of the mountain territory, with an altitude varying between 750 and 1100 m above sea level (a.s.l.).

The nature of the soils, constituting the geological substrate, directly influences the physiography of the area, through the presence of slopes and rocky walls with a good degree of steepness and included among the hydrographic auctions of seasonal streams, which drain limited quantities of water towards the south-eastern sector.

The hydrographic network is characterized by the presence of lower waterways, also defined as incisions or erosive ditches, with south-west towards north-east-oriented feeding the Caperrino stream that physically divides the villages of Castelmezzano and Pietrapertosa. The surfaces present steep slopes in some cases with very high altimetric variations. The area is also affected by abnormal morphological shapes such as ancient and recent niches of detachment, landslides, 'concave–convex' surfaces, as well as the downstream presence of ancient instability phenomena. The presented lithotypes have been subjected to subsequent gravitational elaborations which have determined the granulometric characteristics observed in the outcrop.

From the morphological point of view, there are selective erosion forms developed in the area because of the presence of soils with different characteristics of resistance to the erosion action of exogenous agents. Within the pelitic-arenaceous facies, where the silty-clayey pelitic component prevails, slight slopes and soft morphological configurations are found; in correspondence of the argillitic, marly-clay and arenaceous components, instead, rougher morphologies prevail. From the geomorphological forms observed during the surface relief, it is noted that the area is affected by significant destruction phenomena that could compromise its stability. In particular, the analysed slopes of interest are characterized by reliefs widely modified by the external agents with steep slopes in correspondence of the arenaceous banks. The area under consideration presents sometimes degradation phenomena that affect not only the arenaceous-marly cover but also the arenaceous banks. In fact, especially during the winter months, big rock masses fall from the slope creating problems both to infrastructure lines and to socio-economic activities.

3. The soft computing methods and the analysis neural network (ANN)

The evolutionary genetic algorithms were created from the necessity to simulate the human brain with the available electronic computers, in order to solve the most complicated problems. They have been, since the 1980, successfully applied to many difficult or impossible to treat real world problems. The genetic evolutionary algorithms, together with the fuzzy logic and the artificial neural networks, are part of the soft computing methods in opposition to the traditional ones, hard computing, based on criteria such as precision, determinism and containment of complexity (Tettamanzi et al. 2004). The soft computing is distinguished from the conventional techniques (hard computing) for tolerating the inaccuracy, the uncertainty and the partial truths. It is a discipline in which each methodology completes the others. Among the different Soft Computing (SC) techniques, the artificial neural networks (ANNs) play an essential role. In the field of the automatic learning, an artificial neural network (ANN) is a mathematical model, composed of artificial neurons, inspired by the neural network present in the human brain.

In the recent years, the artificial neural networks (ANNs), frequently applied to the study of complex systems for solving engineering problems of artificial intelligence such as problems related to the various technological fields and, alongside to satellite data, are used to estimate the thematic information at a sub-pixel scale, for the risk analysis (Foody 1997), to classify images, to quantify the land use changes (Kanungo et al. 2006) and finally for the prediction of the environmental dynamics (Follador et al. 2008).

An ANN is a set of simple computational units, called neurons, cells or nodes, linked by a system of connections, the synapses, which work in parallel to realize the input/output transformations. The number of neurons and connections defines the topology of the neural network. An artificial neuron is a mathematical model that simulates a biological neuron (Figure 5).

It is the basic unit of information processing in an ANN and consists of three elements:

- A series of input cells '(X_i) $i = 0 \ldots m$' corresponding to the biological cells that receive the pulse;
- A series of synapses or connections, each of which is characterized by a certain weight '(W_i) $i = 0, \ldots m$'; a positive value is associated with an excitatory synapse and a negative value to an inhibitor synapse. The input information will be integrated through an additive function 'Σ';
- An output cell representing the perceptron response to receive stimuli and is the result of the weighted sum of inputs, limited by a 'g' transfer function (or activation).

A neuron is therefore described mathematically by the following formulas (Mas and Ahlfeld 2007)

$$\begin{gathered} V_k = \sum_{j=1}^{m} w_{kj}x_j + w_{0k} \\ y_k = \varphi(V_k) \end{gathered} \tag{1}$$

where 'w_{0k}' is called bias or activation threshold.

The received inputs are the dendritic analogues in a biological neuron; they are combined, often through a simple weighted sum, to form the internal activation level (Figure 5).

The activation function 'φ (V)' defines the output of a neuron as a linear combination of inputs 'V'. The neuron layout in the network, the number of synapses and nodes define the topology of the ANN. Like biological networks, therefore, the artificial neural ones are composed by a number of processing units operating in parallel. These units are called artificial neurons and they can be broken down into multiple subset of the network, called 'layers'.

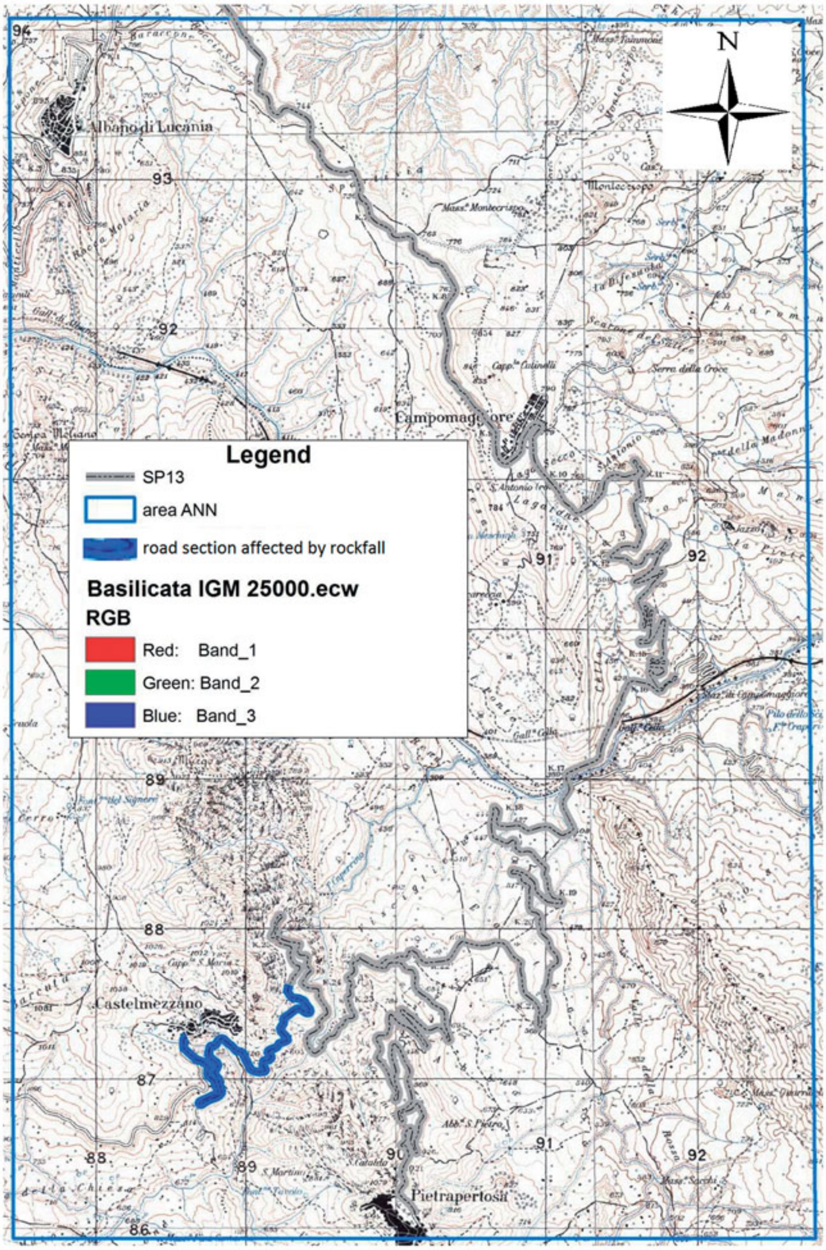

Figure 5. Graphical representation of a network, in which each neuron is described by Eq. (1).

The neurons of each layer can communicate through the weighted connections similar to the biological synapses. Based on the organization of these connections it is possible to distinguish:

- Totally connected networks where each neuron of one layer is connected to each neuron of another layer;
- Partially connected networks, where each neuron of a layer is connected to a particular subset of neurons of another layer;

Or according to the signal flow:

- Probabilistic prediction models for unidirectional entry–exit networks and, for this reason, the model is considered static. The Multi-Layer Perceptron (MLP) is the most used unidirectional class;
- Feedback networks (or recurring networks), where the connections also carry the signal backwards. They are dynamic networks in which a neuron is pushed back directly or passing through intermediate nodes and is reused as input. By means of updating mechanisms, the weights and the bias are modified to the input level. This cycle is repeated until the network converges to a solution.

The most important feature of these systems is to be able to learn mathematical–-statistical models through experience, i.e. by reading experimental data, without determining the mathematical relationships that link the solutions to the problem. The artificial neural network is not programmed, but 'trained' through a learning process based on empirical data.

It is considered a black box model, which contrasts with the 'white box' one because the components inside the system are not known and it cannot be explained logically as it reaches a determined result. There are two different learning/training rules (Kanevski and Maignan 2004): Supervised and Not Supervised.

- *Supervised Learning*: this is used in this work and it is the most common method. It is based on a training phase, building a dataset to 'train' the network, consisting of a series of experimental pairs (real-input, real-output). After reviewing the entire subset of data, the network is updated and the weights modified; this operation aims to minimize the error between output-real and simulated output-network. If the training succeeds, the network learns to recognize the implicit relationship that links the input variables to the output ones and it is able to correctly respond to the stimuli that weren't in the training set;
- *Unsupervised Learning*: only a series of real-input resources are provided to the network, as this method is not accompanied by the 'training' phase (real-output) and it realizes a kind of self-learning. The weights vary according to an a priori defined rule.

3.1. The rockfall susceptibility mapping using the artificial neural network model: the MLP

In this work, a MLP, a three-layer unidirectional input–output model is used to assess the rockfall susceptibility:

- *An Input*: the number of neurons depends on the number of incoming data, such as the thematic maps;
- *An Output*: the number of neurons depends on our goals, such as rockfall susceptibility maps;
- *An hidden layer*: its size and parameterization are chosen on the basis of the results of the network optimization phase.

The MLP is the most important class of a feedforward neural network with supervised training. It consists of a set of inputs (input layer), of computational nodes (hidden layers) and of output nodes (output layer).

A neuron on a 'i' level is always connected to the neuron on the next level '$i + 1$', without any feedback. The MLP learns through a supervised algorithm called 'backpropagation algorithm'. The training of a MLP is based on the following operational steps:

1. Initialization of the weights: random numbers are assigned to synaptic connections and nodes, using a uniform random distribution;
2. Elaboration of the results (forward phase);
3. Error calculation (backpropagation phase).

The forward and backpropagation phases are repeated until a condition completes the learning process. The quality of the results of a trained ANN is verified by calculating the average quadratic error between the network output and the desired real-output. Numerous studies have been carried out to determine the optimum parameters of an ANN (Kavzoglu et al. 2014); this choice depends closely on the examined problem and by the available data.

A last consideration must be made about the data chosen for the network training. The distribution of information and the size of the dataset presented to the ANN, condition the result: to minimize the total error, a network has to process the data that are representative of the real proportions of the studied categories. The data entry order may also affect the network training, since the latest pairs (inputs, outputs) presented to the ANN are more important in updating synaptic weights. In general, this problem is solved by adopting a random introduction order.

3.2. Application of the artificial neural networks (ANN) for the rockfall spatial hazard assessment to the study area

The implemented artificial neural network aims to create a mass movement susceptibility map of the study areas (outputs) subject to rockfalls, by determining: the input parameters, a training phase represented by a portion of the rockfall niches (mass movements already occurred in the past) and a testing phase used to test the network performance. In particular, the methodology used for assessing the rockfall susceptibility for this study involves different phases:

1. Data collection and choice of the input parameters that more influence the slope instability linked to the detachment of rocky blocks from a wall;
2. Analysis of the inventory map of the rockfall phenomenon and detection of rockfall detachment niches with the consequent identification of a representative sample (50% of pixels interested by rockfall – training set and 50% of pixels not interested by the phenomenon – testing set);
3. Realization of the related thematic cartography in 1:25,000 scale through the implementation and the processing of data in GIS environment with relative spatial analysis. Therefore, this phase has resulted in the realization of raster maps of each parameter used as input data in the neural network (GRID format with

resolution 20×20) with the help of the ArcGis 9.3 software. The obtained maps has been then transformed into ASCII format for the implementation in the IDRISI TAIGA software for the realization of the neuronal network;
4. Training phase using the back-propagation algorithm that consists, precisely, in modifying the weight of each connection until the weight set for each parameter is obtained, training the network with reference to 50% of rockfalls niches and 50% of the area not affected by rockfalls until the final map of susceptibility is obtained;
5. Testing phase useful to verify the performance of the model using the remaining 50% of the pixel not affected by mass movements;
6. Development of landslide susceptibility map using the IDRISI TAIGA MLP.

The used module allows to set some parameters such as the value of training pixels for each category, the number of hidden layer (equal to $2n + 1$ where n is the number of parameters used as input data in the neural network) (Hecht-Nielsen 1987), the learning rate that is a positive constant that controls the weights associated with the connections, the momentum factor that prevents the divergence problem during the research of the minimum error value and it is used to accelerate the convergence, the RMSE (the square minimum error) and the number of iteration carried out.

Depending on the entered data, the program returns several maps that are processed in the GIS environment by choosing the definitive susceptibility map, subsequently reclassified into several classes (low, medium, high and very high) based on the inflection points of the cumulative curve of the output provided by the ANN.

The architecture of the neural network for a better understanding has been described in detail below, applying the approach proposed to the considered study area. The area on which the neural network has been 'designed' has an extension of approximately $42\,\text{km}^2$ (Figure 6).

3.2.1. Predisposing factors and relationship with the mass movements

One of the crucial points for the realization of the susceptibility map is the preparation of a spatial database (Lan et al. 2004). The choice of the parameters has been made considering the geomorphological, kinematics and mechanical characteristics of the rockfalls. The basic thematic maps, including the causes or the factors of the mass movement distribution in the study area, have been obtained with the support of the ArcGIS software. Some factors are nominal variables, such as lithology and soil use, the other ones are morphometric (i.e. the slopes) and they have been derived from the Digital Elevation Model (DEM), which has a cell resolution of 20 m $\times$ 20 m (Losasso, Rinaldi et al. 2017).

After rasterizing all the parameters, they have been then converted to ASCII data, that is to say text data that could be processed with any software. The data can be reworked and classified using IDRISI Taiga (Eastman 2009) software for the elaboration of the input maps to be introduced into the MLP module (Multy Layer Perceptron), as shown below, used to generate the artificial neural network. The proposed parameters (Figure 7) to be implemented in an artificial neural network for assessing the rockfall susceptibility in this case study are:

- *Source areas (for each morphological complex)*

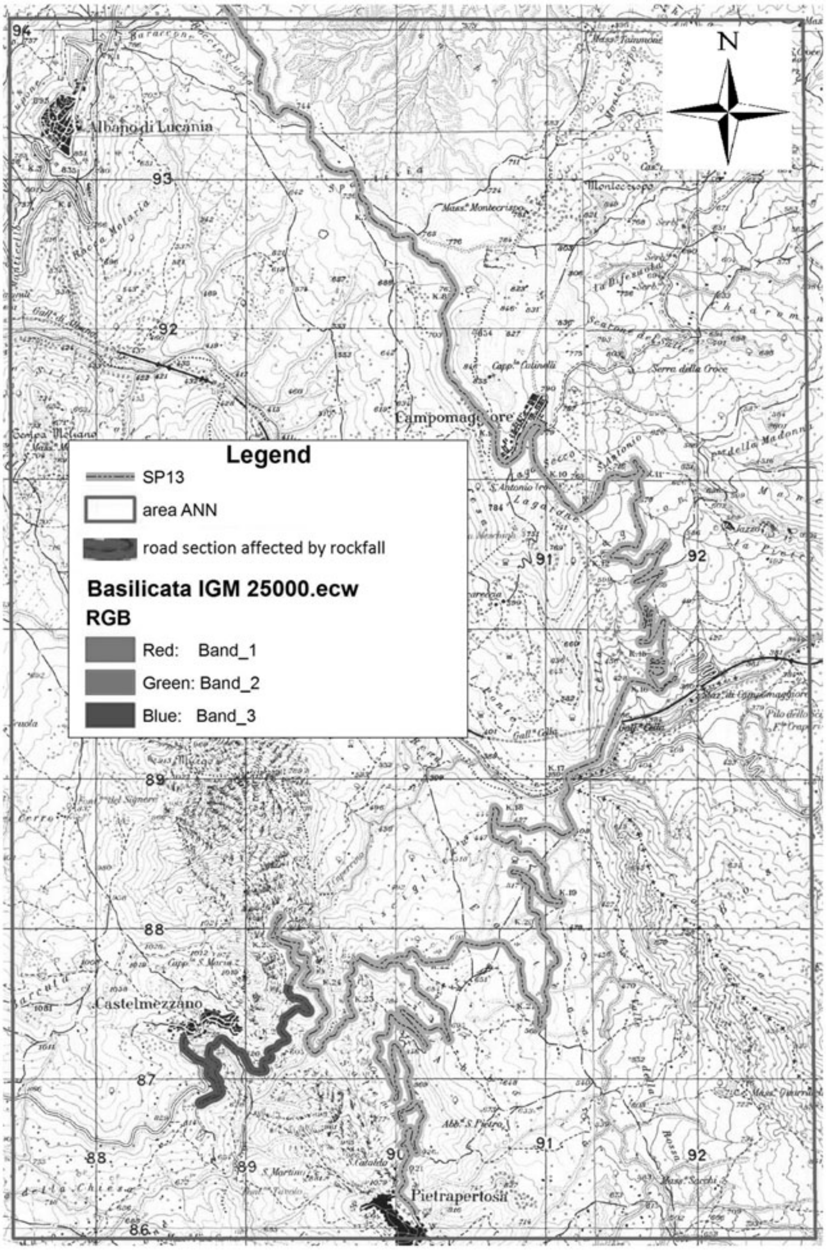

Figure 6. Study area in a portion of the Province of Potenza territory.

- *Lithology*
- *Slope*
- *E-Kin (Kinetic Energy for each morphological complex)*
- DEM
- Rockfall niches
- Land use

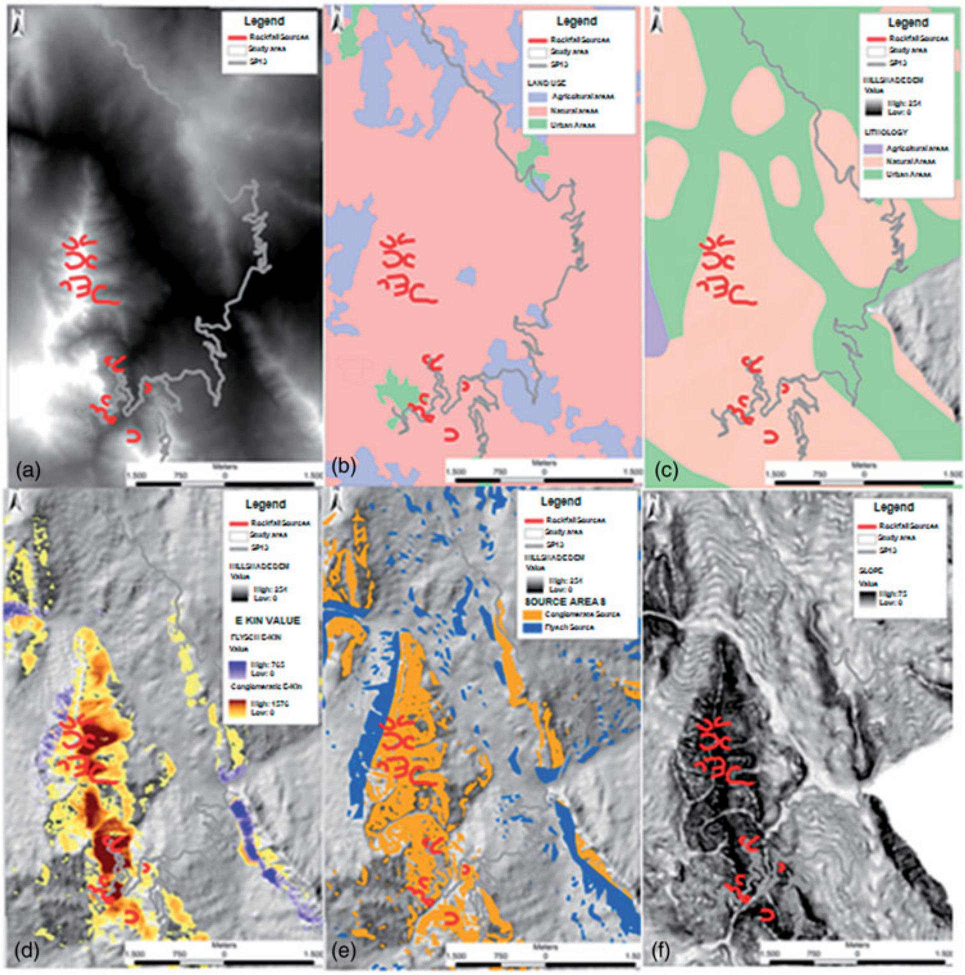

Figure 7. Predisposing input factors: (a) rockfall niches, (b) land use, (c) lythology, (d) kinetic energy value for each morphological unit (Losasso et al., 2016), (e) source areas assessed at a regional scale (Losasso et al., 2016), (f) slope.

Obviously, in different geomorphological conditions, other parameters for the ANN architecture (slope curvature, cracking density, etc.) may be considered.

The input parameters can be both nominal and numeric variables. For this work, starting from previous validated studies (Caniani et al. 2008; Sdao et al. 2013, etc.), it has been decided to use each variable as a sequence of binary numbers to uniform the typology of used data. For this reason, both the numerical and nominal variables have been divided into several classes defined by the influence they exert on the mechanisms of the rockfall phenomena (Losasso et al. 2016).

For the adopted parameters, the following reclassification has been performed (Table 1).

The importance of the analysis of outcropping lithology where rockfalls may occur near a given territory or infrastructure is linked to the dynamics with which the phenomenon occurs, consequently influenced by the type of rock and material that characterizes the analysed territory. Not without reason, the 'lithology' information layer

Table 1. Reclassification of the parameters used for the Analysis Neural Network.

Parameter	Classes	Reclassification
DEM	0	1
	0–200	2
	200–400	3
	400–600	4
	600–800	5
	800–1,000	6
	1,000–1,200	7
	1,200–1,400	8
	1,400–1,600	9
	1,600–1,800	10
Land use	Agricultural areas	1
	Natural areas	2
	Infrastructures and urban areas	3
Lithology	Debris complex	1
	Conglomeratic complex	2
	Flyschoid complex	3
Conglomerate kinetic energy	0–30	1
	30–300	2
	>300	3
Flysch kinetic energy	0–30	1
	30–300	2
	>300	3
Debris kinetic energy	0–30	1
	30–300	2
	>300	3
Rockfall source areas	Potential source areas	1
Slope	0–5	1
	5–10	2
	10–15	3
	15–20	4
	20–25	5
	25–30	6
	30–35	7
	>35	8
Detachement niches		1

is the main predisposing cause of slope instability. The intrinsic characteristics considered in the lithological parameter are the weaving of the materials (nature, shape, size of the grains, nature of the matrix, etc.) and the lithostratigraphic characteristics influencing the stability. The outcropping geological formations in the study area have been grouped in several main complexes based on qualitative assessments (structure, morphology, similar geomechanical behaviour). For the sampling area, three different geological complexes have been recognized: the debris complex, the conglomerate complex and the flyschoid complex.

The DEM (Digital Elevation Models) represent a spatial elevation variation of an area. It has been obtained by digitalizing the level curves from the topographic map 1:25,000 of the Basilicata Region. After the digitalization, the vectorial layer has been converted in raster format with a cell size of 20 m × 20 m. In the study area, the DEM vary between 200 and 1800 meters a.s.l. and it has been subdivided in 8 classes of 200 m.

The slope is another important parameter that belongs to the rockfall predisposing causes of a territory. It represents one of the morphological conditions more affecting the rockfalls. As a consequence, the material accumulated at the base of the slope can be involved in other gravitational phenomena when the morphological conditions allow it.

Further clarification is necessary: the material accumulated at the feet of the slope generally does not participate to the morfoevolutive processes but allows to protect the territory from erosion if the phenomenon occurs near a slope limited to the foot by a stream.

Through the morphological analysis of the digital elevation model it is possible to perform a parametrization of the surface, whose purpose is the numerical description of the continuous form of the surface itself.

Among the topographic attributes most commonly calculated by DEM, the slope plays a fundamental role, used for the evaluation of various parameters linked to numerous geological and geomorphological processes. The slope angle is a parameter that is easily correlated with the movement of a slope, since it is significantly linked to the acting forces. This parameter for the study area varies between 0° and 75° and has been divided into 8 classes.

Land use refers to the different coverage of the land resulting from a 1:50,000 scale cartography with a 44-item legend on 3 hierarchical levels (CORINE LAND COVER 2000). This parameter, used to distinguish the different types of the soil in relation to the type of the use, is useful for identifying the drainage network and the infiltration, since studies have shown that a greater contribution to mass movements is also offered by the exploitation of areas called 'fragile'. In this regard it is important the lack of vegetation that exposes the slopes to the erosive action of rainwater in proportion to their acclivity.

The presence of woods, or the widespread vegetation, in addition to the action of increasing the resistance of the soil by the roots, promotes the interception of significant amounts of rain, avoiding erosion and degradation. This parameter then defines the degree of protection of the soil from the action of atmospheric agents and promotes an important regulating action against the infiltration of surface water, slowing down its time of correction.

The classes identified for this parameter have been reduced to three main classes:

1. Agricultural Areas;
2. Natural Areas;
3. Infrastructures and urban areas.

Another important parameter to take into consideration during the analysis of the rockfall phenomenon is represented by the rockfall niches, that is to say the areas of detachment that can be interpreted as those areas of slope break, often with an arched contour, which can be detected upstream of the collapsed material. Precisely from the morphological point of view, in a mass movement it is possible to distinguish, in addition to the area just described, a riverbed or slope (along which the landslide body moves) and a zone of accumulation.

The preparation of this theme implies a reclassification with assignment of the identifier 1 in the attribute table in correspondence of each rockfall detachment niche. To assess the rockfall susceptibility, it is essential to evaluate the probability of trigger to identify the potential rockfall source areas and consequently the propagation characteristics of the detrital masses, that is to say their path along the slope (probability that the rocky block reaches a certain position in the space). Generally, the triggering cause is linked to an increment in the shear stresses and these mass movements occur mainly in fractured, stratified and/or karstic mass rocks such as limestones, dolomitic limestones, sandstones, conglomerates and magmatic and metamorphic rocks if not altered. Various methods have been proposed in the literature to evaluate both the probability of triggering (Hoek and Bray 1981; Jaboyedoff et al. 1999; Chau et al. 2003; Hantz et al. 2003) and propagation (Desceoudres and Zimmermann 1987; Pfeiffer and Bowen 1989; Evans and Hungr 1993; Agliardi and Crosta 2003), but few attempts have been made to evaluate both (Pierson et al. 1990; Cancelli and Crosta 1993). Also in this case the preparation of this theme implied a reclassification with assignment of the identifier 1 in the attribute table in correspondence with each cell recognized as the source area. The intensity and the recurrence of a phenomenon of falling rocks represent very variable factors that characterize the phenomenon. This process, as it is well known, exhibits itself with high speed, frequency, kinetic energy and mobility, although it involves limited volumes compared with other types of landslides (Evans and Hungr 1993). The rock collapse is predominantly dominated by the free falling motion governed by the gravity. This process is usually described by parabolic trajectories. During the free fall, the potential energy of the boulder is transformed into kinetic energy. Rarely, the main movement of a boulder occurs by pure rolling: usually this tends to be accomplished by a close sequence of rebounds that generate modal height parabolic trajectories (Broili 1973; Azzoni et al. 1991).

3.2.2. Training phase

Training or learning phase is the process that is used to represent the effects of the input parameters and their subgroups on the rockfall probability. The probability of occurrence is determined through a training process, based on a set of examples (training set). In this work, the learning phase of the network is based on a sample of 50% of the pixels falling into the areas represented by the rockfalls (i.e. the niches formed by the blocks already collapsed in the past) and 50% of the pixels in the areas not interested by rockfalls (Figure 8). The training model has been realized with GIS support, starting from a GRID image of the areas represented by the rockfalls niches with 20 m $\times$ 20 m resolution.

The areas interested by the rockfalls are marked with the value 1, while the areas not interested by rockfalls are marked with value 2; the value 0 is assigned to the remaining pixels (NO DATA pixel set). In so doing, the layer used for this study is produced.

3.3. Susceptibility map

The rockfall spatial susceptibility in the study area has been carried out using an artificial neural network tool as widely described above. The used parameters as input data in the network are listed in Figure 9.

So, in light of this, it has been possible to set in the adopted MLP module, a number of hidden layers equal to $2n+1$, that is to say $2 \times 9 + 1 = 19$ hidden layers (Table 1) with a number of carried out simulations equal to 6,996.

In accordance with the data entered, the program allows to obtain several maps drawn up with GIS softwares choosing the definitive susceptibility map.

After the treatment of the images, a map reclassified in Very Low or Nil, Low, Medium, High and Very High susceptibility has been obtained through the

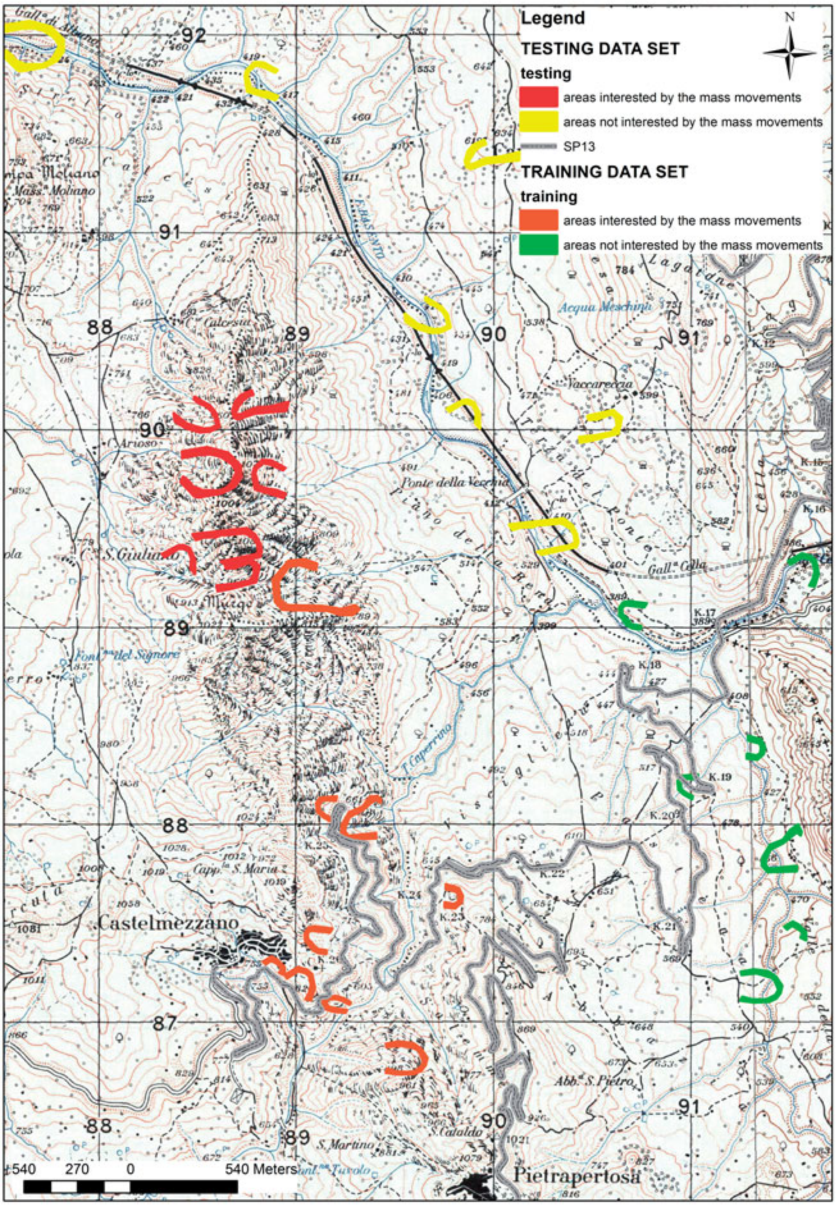

Figure 8. Map representing the 100% of areas interested by mass movements (in blue) and 100% of area not interested by mass movements (in black) divided into training sets and testing sets.

cumulative output distribution provided by the network. The subdivision of the susceptibility ranges in several classes has been carried out according to the natural breaks defined by Jenks (1989) (Jenks Algorithm) (Ruff and Czurda 2008; Falaschi et al. 2009; Nandi and Shakoor 2009).

The *natural breaks* criterion sets the limits between two classes corresponding to discontinuities or ‘jumps’ in the frequency distribution (Figure 10).

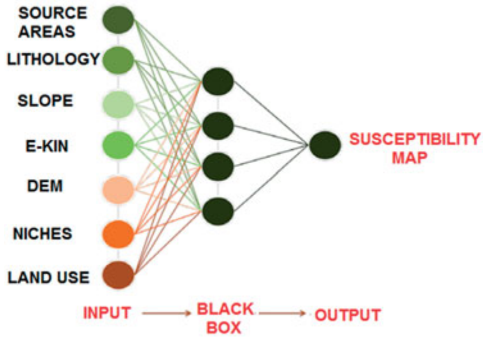

Figure 9. Architecture and running of the artificial neural network for the rockfall susceptibility assessment.

Following the reclassified map according to the five susceptibility classes is presented.

3.3.1. Testing phase

This phase is used to assess the model performance on the data that have not taken part in the training process, and therefore provide a ‘test’ phase for testing a new set of data called ‘testing set’.

This set consists of input and output values unknown by the network and represents, for the study area, approximately the 50% of the remaining pixels containing the rockfall niches. Based on the network response, during the learning iterations, the testing phase demonstrates that most of the mass movements presented as starting data (map of rockfalls detachment niches) fall into the susceptible area considered by the model.

3.4. Validation of the procedure and results

The final phase of the artificial neural network application to the study case is the validation of the procedure that involves two primary goals:

- To analyse where the model is sufficiently accurate, comparing the susceptibility map obtained with the inventory map of the rockfall niches;
- To determine a scheme that can represent instability.

Based on the obtained results, it is possible to decide whether the obtained susceptibility map adequately reflects the initial expectations or if it is necessary to choose an alternative training model and different parameters, repeating the iterative process. For this work, the overlapping of the susceptibility map obtained from the training phase and the map representing the rockfall niches belonging to the testing set has been chosen for the validation of the procedure (Figure 11).

The limited accuracy of the model (Table 2) is certainly due to the preparation of the original shapes concerning the rockfall niches. In particular, after the digitalization of rockfall detachment niches, represented by linear patterns, it has been necessary to turn them into polygons for the creation of the training set and the testing set.

A single niche belonging to the testing set falls into the middle-high hazard class (Figure 12) and therefore this datum is considered a false positive. It consists of 54

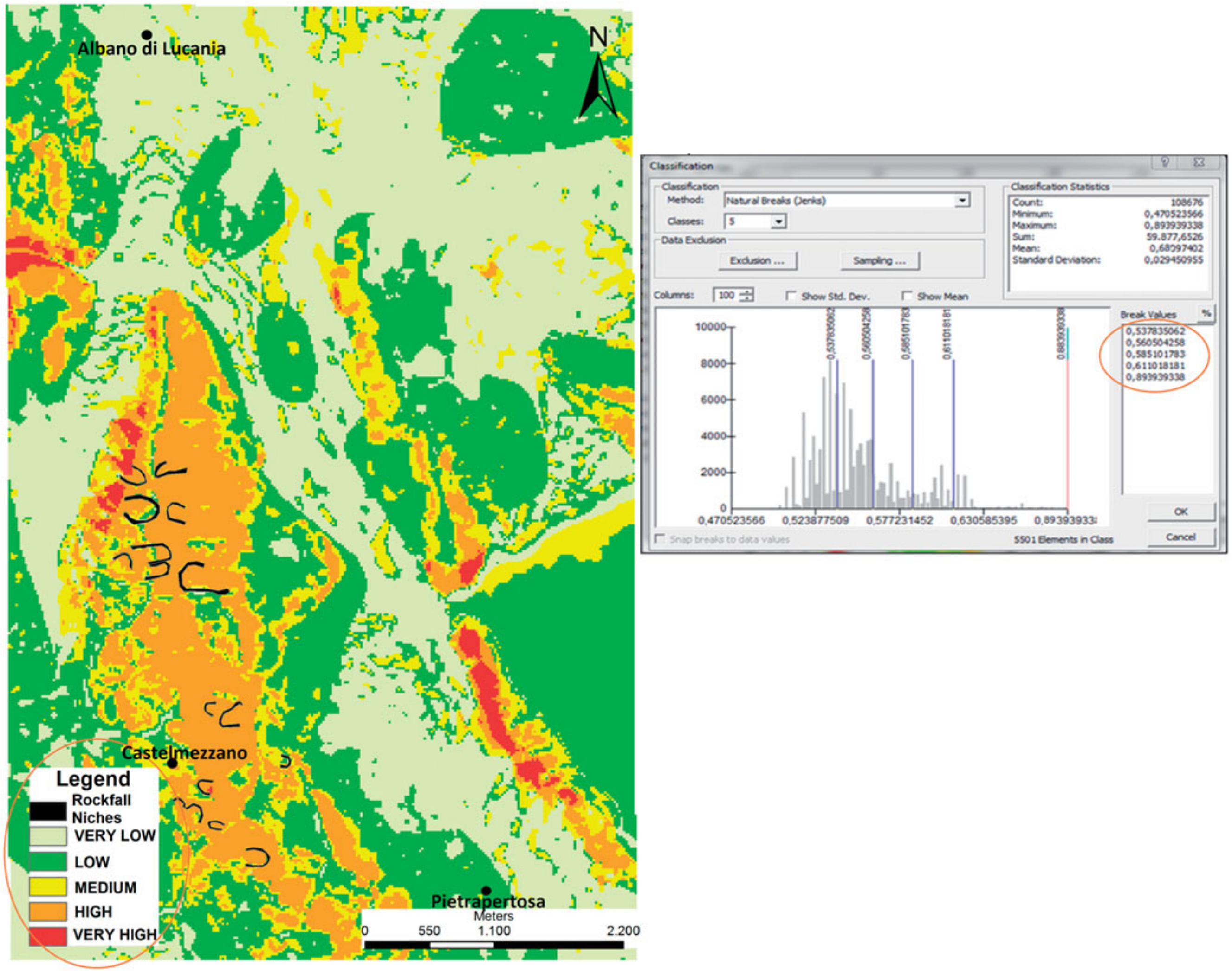

Figure 10. Map of susceptibility (1 represents very low, 2 low, 3 medium, 4 high, and 5 very high of susceptibility level). Example of range of rockfall indices obtained in classes according to the natural breaks defined by Jenks.

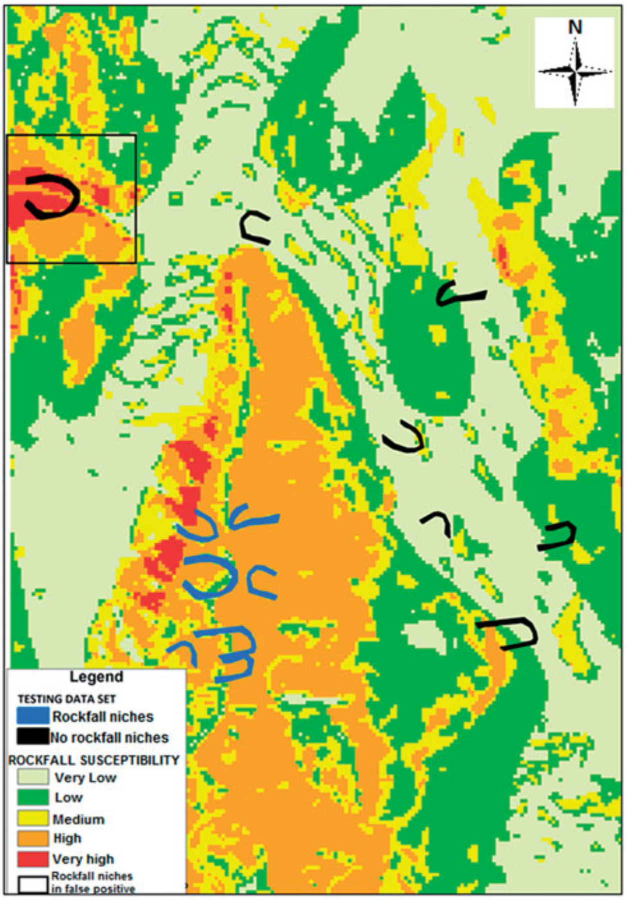

Figure 11. Rockfall susceptibility map compared to the rockfall detachment niches.

Table 2. Assessment of the accuracy of the model.

	Pixels interested by the mass movements	False positive	Pixels not interested by the mass movements	False negative	Tot	Error	Accuracy
Training	153	0	113	0	266	0%	100%
Testing	188	54 pixels not interested by the mass movements fall in the high class	181	0	369	14.63%	85.37%

pixels (this value is very indicative because it depends on the detail scale used for the niches digitalization) which represents a susceptibility estimation error equal to 14.63% (related to false positive in fact there aren't false negative), with subsequent accuracy equal to 85.37%.

Finally, in the considered study area which occupies an area of about 42 km^2 and constituted by 105,000 pixels, 36% fall into the very low hazard class, 36% fall into

Figure 12. Mass movement event occurred in “Margiass” location (14 April 2017).

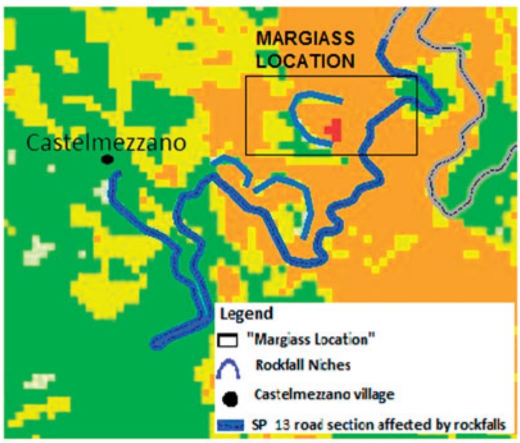

Figure 13. Rockfall susceptibility level in the “Margiass” location (Figure 12).

the low class, 14% fall in the middle class, 12% in the high class and finally 1% falls into the very high class of rockfall spatial hazard.

The rockfall event that occurred in ‘Margiass’ location (Figure 12) near Castelmezzano on 14 April 2017 effectively falls into the very high hazard class (Figure 13).

4. Conclusions

The rockfall susceptibility analysis in the present study has been carried out using an artificial neural network model with a backpropagation learning algorithm and considering the factors that better describe the predisposition of a slope to collapse. In particular, the predisposing factors have been used to obtain a susceptibility map of the study area (lithology, land use, DEM, slope, rockfall niches, rockfall source areas and kinetic energy values for each lithological formation: flyshoid and conglomeratic complex). The training set has been chosen randomly and it is characterized by the 50% of the total pixels interested by the rockfall phenomenon and the 50% of the pixels not interested by the phenomenon in order to train the neural network model.

The number of iteration, carried out by the model in the study area with an extension of about 42 km^2, is equal to 7,000 with an accuracy of about 85%. The final susceptibility map obtained by the elaboration of the input data in the Idrisi Taiga Software has a minimum value equal to 0 and a maximum value equal to 0.89: this means that the pixels with a number close to 0 show low probability of slope instability, while the pixels with a value near to 1 show an high propensity of the slope to collapse. The five susceptibility classes are obtained exporting the data in ASCII format provided by the software IDRISI TAIGA in the GIS environment, finally using the Natural breaks of Jenks that underline the break points according to the identification of the grouped similar values, maximizing the differences between the classes (Federici et al. 2007).

One of the greatest advantages of the Soft Computing methods use such as the artificial neural networks applied in this work is, precisely, the ability of the network to predict a value from a given input data set after the training phase, allowing to solve several problems. Finally, the obtained results underline the fundamental importance of a deepened specialized knowledge of the territorial characteristics predisposing the mass movements for the evaluation of the potential risk conditions, also for planning and civil protection purposes. This characteristic represents the essential and preparatory phase for the hazard and risk assessment. Very important, in fact, is to integrate the subjective experience with the effective use of geo-environmental analysis tools.

A reliable map of susceptibility represents the fundamental step to arrive, through further specialized analyses, to risk and hazard maps. Such documents allow a careful spatial planning and can be usefully integrated in the civil protection systems.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Lucia Losasso http://orcid.org/0000-0002-0361-5110

References

Agliardi F, Crosta G. 2003. High resolution three-dimensional numerical modelling of rockfalls. *Int J Rock Mech Mining Sci*. 40:455–471.

Ayalew L, Yamagishi H. 2005. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. *Geomorphology*. 65:15–31.

Azzoni A, La Barbera G, Mazzà G. 1991. Studio con modello matematico e con sperimentazione in sito del problema di caduta massi. *Bollettino Associazione Mineraria Subalpina Torino*. 28(4):547–573.

Bentivenga M, Capolongo D, Palladino G, Piccarreta M. 2015. Geomorphological map of the area between Craco and Pisticci (Basilicata, Italy). *J Maps*. 11(2):267–277. doi:10.1080/17445647.2014,935501

Broili L. 1973. In situ tests for the study of rockfall. *Geol Appl Idrogeol*. 8(1):105–111.
Cancelli A, Crosta G. 1993. Rockfall hazard and risk mapping. In: Novosad S, Wagner P, editors. Proceeding of the 7th International Conference and Field Workshop on Landslide in Czech and Slovak Republics; Balkema, Rotterdam.
Caniani D, Pascale S, Sdao F, Sole A. 2008. Neural networks and landslide susceptibility: a case study of the urban area of Potenza. *Nat Hazards*. 45:55–72.
Carrara A. 1983. Multivariate models for landslide hazard evaluation. *Math Geol*. 15:403–426.
Carrara A, Cardinali M, Detti R, Guzzetti F, Pasqui V, Reichenbach P. 1991. GIS techniques and statistical models in evaluating landslide hazard. *Earth Surf Processes Landforms*. 16:427–445.
Carrara A, Pugliese Carratelli E, Merenda L. 1977. Computer-based data bank and statistical analysis of slope instability phenomena. *Geomorphology*. 21(2):187–222.
Chau KT, Wong RH C, Lui J, Lee CF. 2003. Rockfall hazard analysis for HongKong based on rockfall inventory. *Rock Mech Rock Eng* 36(5):383–408.
Chung CF, Fabbri AG. 1999. Probabilistic prediction models for landslide hazard mapping. *Photogrammetric Engineering and Remote Sensing*. 65(12):1389–1399.
Desceoudres F, Zimmermann T. 1987. Three-dimensional dynamic calculation of rockfalls. In: Proceedings 6th International Congress of Rock Mechanics; Montreal, Canada. p 337–342.
Eastman JR. 2009. *Idrisi Taiga, guide to GIS and image processing, user's guide (Ver. 15)*. Worcester (MA), USA: Press Clark University; 328 p.
Ercanoglu M, Temiz F A. 2011. Application of logistic regression and fuzzy operators to landslide susceptibility assessment in Azdavay (Kastamonu, Turkey). *Environ Earth Sci*. 64:949–964.
Evans S, Hungr O. 1993. The assessment of rockfall hazard at the base of talus slopes. *Can Geotech J*. 30:620–636, doi:10.1139/t93-054
Falaschi F, Giacomelli F, Federici PR, Puccinelli A, D'Amato Avanzi G, Pochini A, Ribolini A. 2009. Logistic regression versus artificial neural networks: landslide susceptibility evaluation in a sample area of the Serchio River valley, Italy. *Nat Hazards*. 50:551–569.
Federici PR, Puccinelli A, Cantarelli E, Casarosa N, D'Amato Avanzi G, Falaschi F, Giannecchini R, Pochini A, Ribolini A, Bottai M, et al. 2007. Multidisciplinary investigations in evaluating landslide susceptibility – an example in the Serchio River valley (Italy). *Quat Int*. 171(172):52–63.
Follador M, Villa N, Paegelow M. 2008. Tropical deforestation modelling: comparative analysis of different predictive approaches. The case study of Peten, Guatemala. In: Paegelow M, Camacho Olmedo MT, editors. *Modelling environmental dynamics*. Berlin: Heidelberg: Springer; p. 77–107.
Foody GM. 1997. Fully fuzzy supervised classification of land cover from remotely sensed imagery with an artificial neural network. *Neural Comput Appl*. 5:238–247.
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.
Hantz D, Vengeon JM, Dussauge C. 2003. An historical, geomechanical and probabilistic approach to rockfall hazard assessment. *Nat Hazards Earth Syst Sci*. 3:693–701.
Hecht-Nielsen R. 1987. Kolmogorov mapping neural network existence theorem. *IEEE First International Conference on Neural Networks*. 3:11–14.
Hoek E, Bray J. 1981. *Rock slope engineering*. 3rd ed. London, United Kingdom: Institute of Mining and Metallurgy; p. 211–219.
Jaboyedoff M, Baillifard F, Marro C, Philippossian F, Rouiller J-D. 1999. Detection of rock instabilities: matterock methodology. In: Masuya H, Labiouse V, editors. Joint *Japan-Swiss Seminar on Impact Load by Rockfalls and Design of Protection Structures;* 4–7 October 1999; Kanazawa, Japan. p. 13–19.
Jenks GF. 1989. Geographic logic in line generalization. *Cartographica*. 26:27–42.
Kanevski M, Maignan M. 2004. *Analysis and modelling of spatial environmental data*. Lausanne Switzerland: EPFL Press.

Kanungo DP, Arora MK, Sarkar S, Gupta RP. 2006. A comparative study of conventional, ANN black box, fuzzy and combined neural and fuzzy weighting procedures for landslide susceptibility zonation in Darjeeling Himalayas. *Eng Geol*. 85:347–366.

Kavzoglu T, Sahin EK, Colkesen I. 2014. Landslide susceptibility mapping using GISbased multi-criteria decision analysis, support vector machines, and logistic regression. *Landslides*. 11(3):425–439.

Lan HX, Zhou CH, Wang LJ, Zhang HY, Li RH. 2004. Landslide hazard spatial analysis and prediction using GIS in the Xiaojiang watershed, Yunnan, China. *Eng Geol*. 76:109–128.

Lee S, Hwang J, Park I. 2012. Application of data-driven evidential belief functions to landslide susceptibility mapping in Jinbu, Korea. *Catena*. 100:15–30.

Losasso L, Derron M-H, Horton P, Jaboyedoff M, Sdao F. 2016. Definition and mapping of potential rockfall source and propagation areas at a regional scale in Basilicata region (Southern Italy) *Rendiconti Online Della Società Geologica Italiana*. 41:175–178. DOI:10.3301/ROL.2016.12

Losasso L, Jaboyedoff M, Sdao F. 2017. Potential rock fall source areas identification and rock fall propagation in the province of Potenza territory using an empirically distributed approach. *Landslides*. 14:1593–1602.

Losasso L, Pascale S, Sdao F. 2017. Landslides risk assessment along roads: the transportation corridors of the "Dolomiti Lucane" (Basilicata). In: Mikoš M, et al., editors. Advancing Culture of Living with Landslides. Proceedings of WLF 4, Vol. 4; p. 661–667. DOI: 10.1007/978-3-319-53485-5_76.

Losasso L, Rinaldi C, Alberico D, Sdao F. 2017. Landslide risk analysis along strategic touristic roads in Basilicata (Southern Italy) using the modified RHRS 2.0 method. In: Proceedings of 17th International Conference on Computational Science and Its Applications, ICCSA 2017, July 3-6, 2017; Trieste, Italy.

Mas DML, Ahlfeld DP. 2007. Comparing artificial neural networks and regression models for predicting faecal coliform concentrations. *Hydrol Sci J*. 52:713–731. DOI:10.1623/hysj.52.4.713

Nandi A, Shakoor A. 2009. A GIS-based landslide susceptibility evaluation using bivariate and multivariate statistical analyses. *Eng Geol*. 110:11–20. DOI: 10.1016/j.enggeo. 2009.10.001

Pfeiffer TJ, Bowen T. 1989. Computer simulation of rockfall. *Bulletin of the Association of Engineering Geologists*. 26(1):135–146.

Pierson LA, Davis SA, Van Vickle R. 1990. Rockfall hazard rating system – implementation manual, Federal Highway Administration (FHWA), Report FHWA-OR-EG-90-01. FHWA, U.S. Dep. of Transp.

Ruff M, Czurda K. 2008. Landslide susceptibility analysis with a heuristic approach in the Eastern Alps (Vorarlberg, Austria). *Geomorphology*. 94:314–324.

Sdao F, Lioi DS, Pascale S, Caniani D, Mancini IM. 2013. Landslide susceptibility assessment by using a neuro-fuzzy model: a case study in the Rupestrian heritage rich area of Matera. *Nat Hazards Earth Syst Sci*. 13:395–407.

Sdao F, Sole A. 2013. Scientific-Technical advice: "Valutazione del rischio idrogeologico combinato (frane ed alluvioni) e del relativo grado di esposizione della rete stradale della Provincia di Potenza" – Province of Potenza.

Tettamanzi A, Sammartino L, Simonov M, Soroldoni M, Beretta M. 2004. Learning environment for life-time value calculation of customers in insurance domain. In: Deb K, et al., editor. Proceedings of the Genetic and Evolutionary Computation Congress (GECCO 2004), S. Francisco, 26–30 Giugno 2004, p. 1251–1262.

Van Westen CJ. 1993. *Application of geographic information systems to landslide hazard zonation*. Enschede, The Netherlands: Int. Ins. for Aerospace and Earth Res. Surv; p. 245. ITC publ. no. 15.

van Westen CJ, Terlien MJT. 1996. An approach towards deterministic landslide hazard analysis in GIS. A case study from Manizales (Colombia). *Earth Surf Process Landf*. 21(9): 853–868