Strategies investigation in using artificial neural network for landslide susceptibility mapping: application to a Sicilian catchment

Elisa Arnone, Antonio Francipane, Leonardo V. Noto, Antonino Scarbaci and Goffredo La Loggia

ABSTRACT

Susceptibility assessment of areas prone to landsliding remains one of the most useful approaches in landslide hazard analysis. The key point of such analysis is the correlation between the physical phenomenon and its triggering factors based on past observations. Many methods have been developed in the scientific literature to capture and model this correlation, usually within a geographic information system (GIS) framework. Among these, the use of neural networks, in particular the multi-layer perceptron (MLP) networks, has provided successful results. A successful application of the MLP method to a basin area requires the definition of different model strategies, such as the sample selection for the training phase or the design of the network structure. The present study investigates the effects of these strategies on the development of landslide susceptibility maps by applying different model configurations to a small basin located in northeastern Sicily (Italy), where a number of historical slope failure events have been documented over the years. Model performances and their comparison are evaluated using specific metrics.

Key words | artificial neural network, GIS, landslide, susceptibility mapping

Elisa Arnone (corresponding author)
Antonio Francipane
Leonardo V. Noto
Antonino Scarbaci
Goffredo La Loggia
Dipartimento di Ingegneria Civile, Ambientale, Aerospaziale, dei Materiali,
Università degli Studi di Palermo,
Palermo,
Italy
E-mail: *elisa.arnone@unipa.it*

INTRODUCTION

Every year landslide events hit various areas throughout the world, often causing severe economic and social damage, thus making the field of landslide prevention a current issue in land management.

The availability of innovative technologies and the growing need for an efficient management of natural hazards have recently led to the development of new methods for landslide prevention analysis, following on from the classical deterministic ones (Arnone *et al.* 2011). In general, the methods based on the recognition of landslide-prone terrains are traditionally considered one of the most useful approaches (Hansen 1984), whose ultimate goal is to define susceptibility maps.

The term 'susceptibility' commonly refers to the probability of a landslide occurrence over a region, based on empirical or modeled relationships between historical events and the so-called landslide-inducing or triggering factors (Varnes & IAEG 1984). It follows that the estimation of susceptibility results in a spatial correlation analysis between the triggering factors and landslides occurrences.

The scientific literature of the last two decades has highlighted how statistical and geographic information system (GIS)-based methodologies have been particularly successful in the definition of susceptibility maps. Among the statistical methods, both the bivariate and the multivariate approaches have been widely used in many susceptibility analyses. By using the bivariate statistical methods (Lee & Pradhan 2007; Lepore *et al.* 2012), the contribution of each landslide-inducing factor to slope failure susceptibility is evaluated individually, overlooking the possibility that the different factors may have a mutual relationship. In multivariate statistical methods, all factors are analyzed using multiple

doi: 10.2166/hydro.2013.191

regression techniques. In particular, the logistic regression model (Hosmer & Lemeshow 2000) is well suited to analyze the presence/absence of a dependent variable (Carrara *et al.* 1991; Guzzetti *et al.* 1999; Lee *et al.* 2004b; Lee & Pradhan 2007; Lepore *et al.* 2012), representing one of the most applied methods. The multivariate statistical methods offer the advantages of significantly reducing the number of factors to be analyzed, allowing the identification of additional variables that more influence the dependent variable.

However, these statistical approaches are based on some assumptions that often do not match the properties of the distribution of landslides factors. Also, because of these limitations, recently, new techniques have been explored that favor a data-driven modeling approach.

Among the data-driven methods, the artificial neural network (ANN) is the model that overcomes most of the above-mentioned limits. This model has the ability to handle a large amount of information and to learn complex model functions from examples, i.e., by 'training', using sets of input and output data (Giustolisi & Savic 2006).

The use of ANNs for landslide susceptibility analysis has seen an increasing number of successful applications in recent years (Lee *et al.* 2004b; Ermini *et al.* 2005; Yesilnacar & Topal 2005; Caniani *et al.* 2008; Melchiorre *et al.* 2008; Nefeslioglu *et al.* 2008; Bai *et al.* 2009; Falaschi *et al.* 2009; Pradhan & Lee 2009; Doglioni *et al.* 2012). An interesting combined use of the logistic regression model and ANN method has been proposed by Lee *et al.* (2004b), who used a probabilistic method to calculate the rating of the relative importance among the classes of each triggering factor, and an ANN method to calculate the weight of the relative importance of each triggering factor. Nefeslioglu *et al.* (2008) showed that ANNs provide a more optimistic evaluation of landslide susceptibility than logistic regression analysis, whereas Melchiorre *et al.* (2008) improved the predictive capability and robustness of ANNs by introducing a cluster analysis.

One of the most critical issues in applying an ANN for landslide analysis is the selection of the proper dataset to use to train the model structure. This issue is commonly ignored and not well stated in scientific works (Nefeslioglu *et al.* 2008). To our knowledge, a clear and universal criterion has not been provided yet. Among a few studies, Yilmaz (2010) makes an interesting contribution to this topic by analyzing the effects of different sampling strategies (scarp, seed cells, point) on the definition of landslide susceptibility maps, demonstrating that scarp sampling strategy performs better than the point strategy, which is commonly used in many works. Moreover, Tian *et al.* (2008) and Lee *et al.* (2010) show how the available resolution of the landslide inventory map can affect the choice of the input map resolution, and the use of an input map with finer resolution does not necessarily increase the accuracy of landslide susceptibility maps.

Starting from the work described in Arnone *et al.* (2012), this study proposes the use of the ANN-based methodology for the landslide susceptibility mapping of a small Sicilian catchment (Italy), where a number of historical events have been documented over the years. In particular, it will investigate how sampling strategies and the design of the neural network structure can improve the resulting maps. The results from the comparison of these different strategies may provide important indications for future analyses.

METHODS

Artificial neural networks

The use of ANNs is a valid alternative to the classical statistical methods when a multivariate approach is needed, such as in a landslide susceptibility analysis. An ANN is a collection of basic units, called neurons, computing a nonlinear function of their input and able to perform pattern recognition and classification (Haykin 1994). The main characteristic of ANNs consists of the ability to derive rules from multivariate data after self-learning the reality and then to reproduce predictive patterns. Every input has an assigned weight that determines the influence of this input on the overall output of the node.

Among the different types of ANNs, the multi-layer perceptron (MLP) network is currently considered the most well-used type (Ermini *et al.* 2005), especially in landslide susceptibility analysis (Lee *et al.* 2003), and thus it was chosen in this application.

The architecture of the MLP consists of a number of neurons connected by weighted links and placed into different layers: an input layer, with a number of neurons equal to the number of the independent variables, an output layer,

with a number of neurons equal to the number of the dependent variables, and ultimately one or more so-called hidden layers. The MLP is a feedforward ANN, where an arbitrary input vector is propagated forward through the network without feedbacks (Hines 1997). Layers are fully connected to each other, i.e., each neuron on a layer is connected to every neuron on the following layer. The hidden neuron layers make a linear combination of input signals, convert them through a generally nonlinear function (activation function), to then transfer the information to the next layer. The number of hidden layers and the number of nodes in a hidden layer which are required for a particular classification problem are not easy to deduce (Lee *et al.* 2003). MLP networks with one hidden layer are commonly used in landslide susceptibility modeling and they are considered to provide enough complexity to accurately simulate the nonlinear properties of the landslide process (Lippmann 1987).

Neural networks are also called 'machine learning algorithms', because the changing of their connection weights causes the network to learn the solution to a problem. The strength of a connection between the neurons is stored as a weight-value for each specific connection. The system learns new knowledge by adjusting these connection weights. The key instrument that allows the network to learn the dynamics of a particular phenomenon is called training phase. During the training phase a set of known input–output couples are fed to the network and the weights are updated by following some pre-determined learning rule, so that the resulting output vector of the net is almost equal to the target vector. Weights are updated to minimize a cost function E and the distance between the target and the actual output vector. The learning ability of a neural network is determined by its architecture and by the algorithm chosen for training. The 'error back propagation' (Werbos 1994) is the most widely used learning algorithm to perform a gradient descent along the cost surface of ANNs (Lee *et al.* 2003), and it is also used in this study. Information about errors is filtered back through the system and is used to adjust the connections between the layers, improving performance.

Assessment of model performance

Model performance of ANNs is evaluated by means of the area under the curve (AUC) of the receiving operating characteristic (ROC) curves (Fawcett 2006). The AUC ranges from 0 to 1 and gives a measure of the model's ability to discriminate between elements that experienced the outcome of interest (landslide occurrence) versus those which did not. An AUC value of 1 indicates a perfect fit of the model to reality, while a value of 0.5 represents a fit indistinguishable from random occurrence. The measure of the model's ability in identifying correctly the elements which experience the outcome of interest (true positive rate) is called 'sensitivity', while the measure of the model's ability in identifying correctly the elements that do not experience the outcome of interest (true negative rate) is known as 'specificity'. The ROC curve is built plotting the sensitivity, versus 1-specificity, over all the possible cutoffs, i.e., those threshold values separating the two opposite states (in this case landslide and no landslide).

CASE STUDY: THE TIMETO CATCHMENT

The Timeto catchment is located in northeastern Sicily, between the Peloritani Mountains and the Tyrrhenian coast, within the Messina district (Figure 1(a)). The catchment is approximately 95 km^2 in size and its elevation ranges between 0 and 1,350 m a.s.l. (Figure 1(a)). The highest number of landslides and the largest area of soil removed by landslides in Sicily characterize this area. Because of its well-documented slope failure history, this basin has been used in the evaluation of landslide susceptibility maps.

The basin morphology consists of a mainly flat coastal region with low slopes toward the sea, and a mountain region with a rugged morphology, narrow valleys, and very steep hillslopes. The region between the coast and the mountains is characterized by hills with variables slopes.

The climate of the catchment is typical of the Mediterranean area. Mean annual precipitation (MAP) ranges between 700 and 800 mm in the coastal and hill regions, and 800 and 1,000 mm in the mountain region (Di Piazza *et al.* 2011). The mean annual temperature is 18 °C with the average monthly maximum temperatures of 30 °C in July, and the average monthly minimum temperatures of 4.5 °C in February.

The runoff regime of the Timeto river is ephemeral like many other rivers in northeastern Sicily, with low-flow or

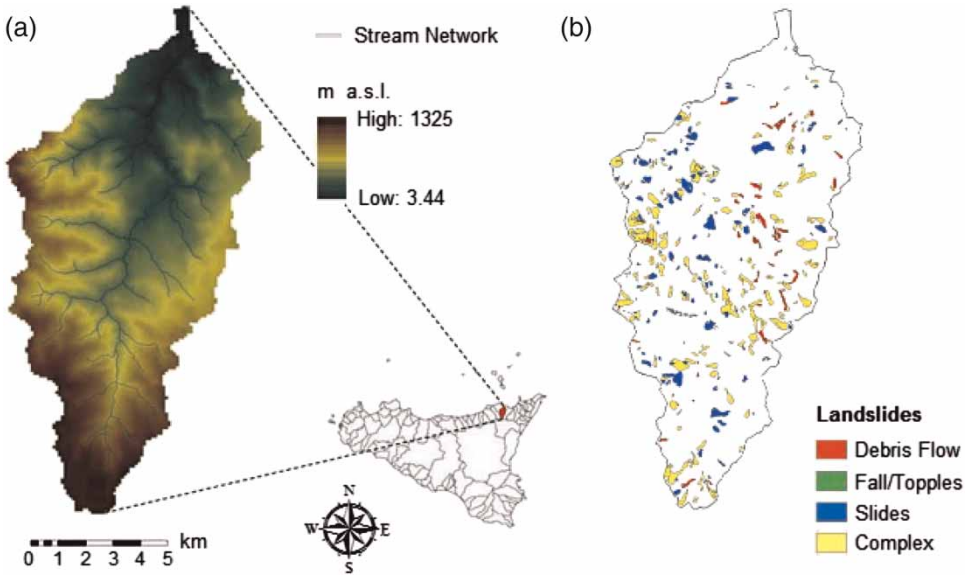

Figure 1 | Study area: (a) digital elevation model and (b) landslide inventory of the Timeto catchment.

null discharges during the dry season and high-flow discharges during the autumn and the winter (Viola *et al.* 2011).

Landslide inventory

The landslides inventory has been derived from the Carta Inventario delle Frane, made by the Assessorato Territorio ed Ambiente della regione Sicilia in 2006 (P.A.I. 2006). The map reports around 700 landslides inside the Timeto catchment and classifies them according to the Varnes' classification (Varnes & IAEG 1984). Starting from this map, the landslide inventory for the Timeto catchment has been created (Figure 1(b)) including only four landslide classes (flows, falls/topless, slides, and complex landslides). The resulting inventory map describes the landslide events by polygons, according to the scarp sampling strategy of Yilmaz (2010). The average value of landslide areas is about 2.1 ha and more than 85% of landslide polygons have an area greater than 1 ha; finally, almost 10% of the basin area has experienced slope failure.

Landslide-inducing factors

In any method used for landslide susceptibility analysis, the initial choice of landslide-inducing factors can be considered the most subjective step. However, past experiences, the knowledge of the site, and several examples from the literature can be useful in identifying the factors that may mostly affect slope failure. In this study, the following factors have been chosen, based also on the data availability:

- *Morphometric parameters*, such as aspect (Figure 2(a)), slope angle (Figure 2(b)), profile curvature (Figure 2(c)), and distance from the stream network (Figure 2(d)).
- *Geological characteristics*, such as lithology and distance from the faults.
- *Hydro-climatic characteristics*, such as the MAP, and the hydrological parameters a and n of the rainfall depth–duration curve, expressed as $h_{d,t} = a_T d^{n_T}$; $h_{d,T}$ is the hydraulic head for given duration d and return period T of rainfall. The MAP data come from Di Piazza *et al.* (2011), while the parameters a and n were derived from Lo Conti *et al.* (2007), and allow one to take into account the effect of very intense rainfall on landslide susceptibility.
- *Other spatial information*, such as land use, derived from the Corine Land Cover (APAT 2005) map (Figure 2(e)), soil types map, derived from the Sicilian soil map (Figure 2(f)) (Fierotti & Ballatore 1988), and distance from the roads, considering how road construction often implies an important disruption of the original slope.

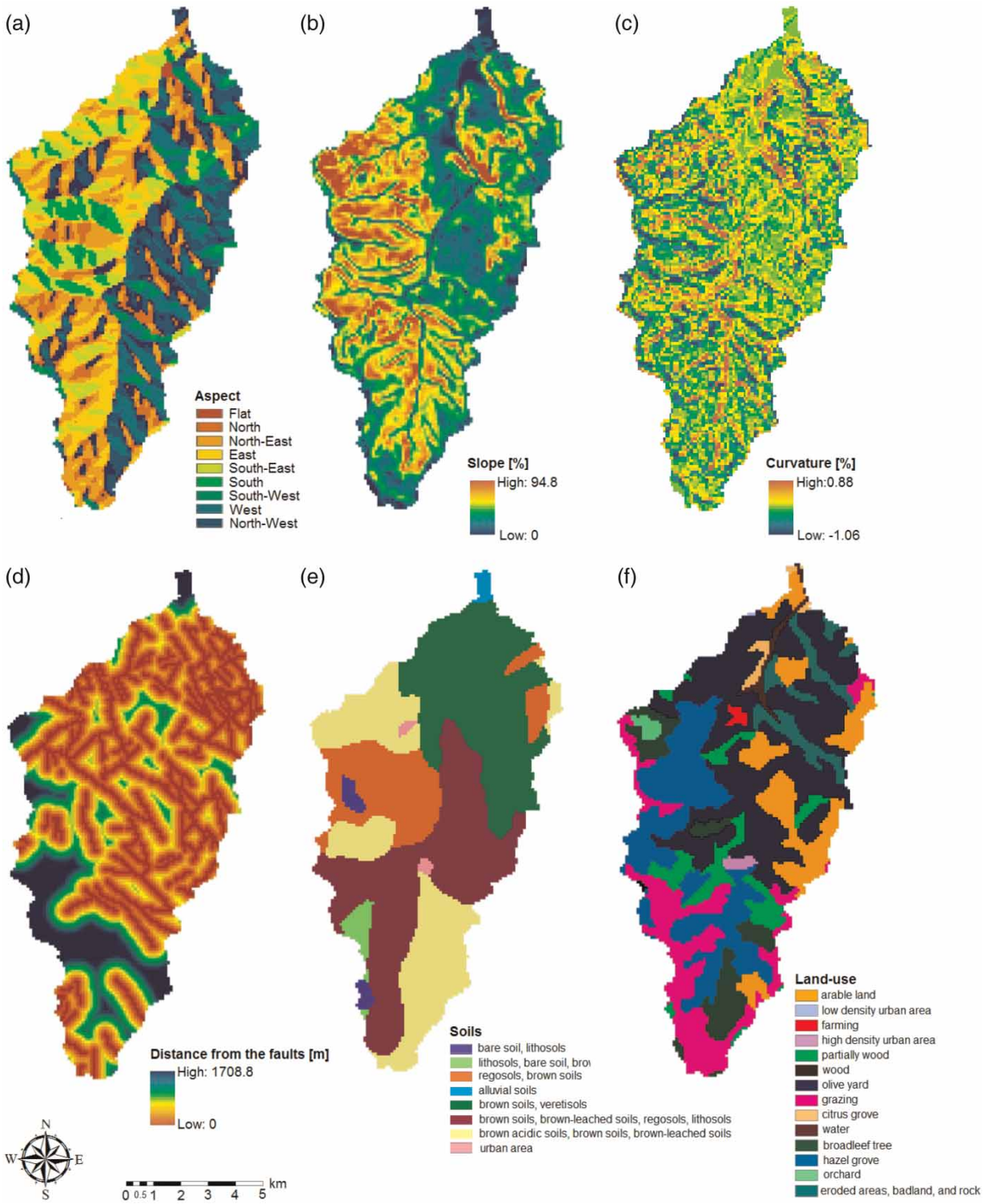

Figure 2 | Some characteristics of the Timeto catchment used as landslide inducing factors: (a) aspect, (b) slope, (c) curvature, (d) distance from the faults, (e) soils, and (f) land use maps.

Similarly to the landslide inventory, all these factors have been used to build a spatial database in the form of maps (see the section below). Each map represents a possible inducing factor for rainfall-induced landslides on the Timeto catchment.

GIS database

All the data listed in the previous two paragraphs (landslide location, morphology, land use, soils, geology, climate forcing, etc.) have been used to create the spatial database used in the analysis. The spatial database was built using ArcGIS/ArcINFO software.

The grid maps of the primary morphometric parameters (Wilson & Gallant 2000), such as slope, profile curvature, and aspect, were derived from a 100-m resolution digital elevation model (DEM) (resampled by the original 30-m resolution DEM produced by 'Regione Sicilia'), by means of spatial analysis techniques.

Other data, such as the landslide inventory map, the 1:250,000 scale soil map, the 1:50,000 scale geology map, and the 1:100,000 scale land use map, were digitized into vector structures (shapefile format) and then converted into a grid structure using the same spatial extent and resolution of the 100-m DEM. Finally, all the grid maps have the same resolution.

The choice of a 100-m DEM was induced and supported by various reasons, particularly: (1) the pixel size is comparable with most of the real landslide areas (1 ha) sampled with the scarp strategy; (2) according to Tian *et al*. (2008) and Lee *et al*. (2010) a resolution finer than the landslide scale it is not preferable; (3) the scale of some of the used input map (geology, pedology, land use) is not very accurate, suggesting the use of a coarser resolution, according to Lee *et al*. (2004a). Finally, the main target of the work is to investigate and compare different model strategies, therefore the use of the same accuracy is more significant than the level of accuracy itself.

The landslide map (Figure 1(b)) was converted and reclassified into a binary grid map with 1 for landslide cells and 0 for no landslide cells.

Over the entire study area of 100 m ×100 m, 9,593 cells, 928 are landslide cells, corresponding to about 9.7% of the basin area, while the remaining 8,665 cells are no landslide cells.

All the grid maps were finally exported to the Esri ASCII grid format, to use them within the Neural Network toolbox of MATLAB software (Mathworks), here chosen to design and apply the MLP network.

Application of artificial neural network model

As described in the section 'Artificial neural network', the phases involved in the development of the MLP network are various and need to be carefully defined in order to build up the most successful model. Such steps are identified as follows: (1) data selection for the network training, (2) MLP design (input, hidden and output layers, and number of nodes), (3) network training (choosing activation and transfer functions), and (4) classification.

(1) *Data selection for the network training*

The selection of a proper dataset for the network training is far from being obvious and represents one of the most critical steps, although not many works on landslide susceptibility have noted this. It is common to select a random subset of cells corresponding to a portion of the entire database (Lee *et al*. 2003; Ermini *et al*. 2005), which includes both landslide and no landslide cells. However, in most studies, the selected subset rarely maintains the actual ratio between landslide and no landslide areas, and the percentage of landslide cells in the subset is often increased with respect to the total area. In particular, Lee *et al*. (2003) randomly selected a fixed number of cells from each of the two classes (landslide and no landslide), setting a ratio between landslide and no landslide cells equal to 1:1, despite the recorded ratio relative to the entire area being equal to 1:56. They analyzed three different sizes of cell number, 200, 400, and 600 pixels/class, respectively, which corresponded to percentages of 1.7, 3.4, and 5.1% of the landslide cells and 0.03, 0.06, and 0.09% of the no landslide cells. They claimed that the number of training locations had little influence on the analysis. In Ermini *et al*.'s (2005) work, the training phase was executed on a subset corresponding to one-third of the entire database, randomly selected from the whole dataset. Caniani *et al*. (2008) randomly selected 32% of the entire landslide database (which in turn corresponds to 34% of the entire basin). Melchiorre *et al*. (2008) adopted cluster analysis

technique for the selection of training data, by using the early stopping technique (Caruana *et al*. 2001).

In our application, three different subsets were defined in order to pursue the best configuration. Both the methodological (one subset) and the random (two subsets) criteria have been applied.

In the case of the methodological criterion, 10 kernels of 15×15 pixels were placed over the entire basin, selecting the pixels (landslide and no landslide) falling therein. The kernels were especially placed in such a way to ensure that landslide pixels were distributed over the basin, avoiding overlaps. In this way subset 1 was obtained, which counts 399 landslide cells and 1,741 no landslide cells and includes 43% of landslide database and 20.1% of the no landslide area, respectively (Table 1).

In the random case, we randomly selected 50% of landslides cells (464 cells) while the no landslide cells were randomly selected in numbers reflecting two different ratios between landslides and no landslide cells, equal to 1:1 (subset 2), according to Lee *et al*. (2003), and 1:2 (subset 3) (see Table 1 for details).

(2) *MLP design*

The definition of the MLP structure requires the definition of input, hidden and output layers, and number of nodes for each layer.

The structure of the input vector depends on (i) number and type (continuous or categorical) of triggering factors and (ii) methodology used in representing the data. Here the methodology presented in Chung *et al*. (1995) and Ermini *et al*. (2005) has been used, representing each variable as a sequence of binary numbers. First of all, the approach requires categorization of each landslide factor in classes; in the case of categorical factor (such as land use, pedology, etc.), the original classes were kept; in the case of continuous variable (e.g., slope, distance from road, etc.), a quantile-based classification was used. The number of classes for each variable is shown in Table 2, for a total number of 73 classes. At each computational cell (i.e., each cell of the basin), the input vector is represented by a string of binary values indicating whether the cell belongs (1) or not (0) to each of the 73 classes. As an example, Figure 3(a) shows the binary string of a given cell characterized by an aspect value falling into class 9, a curvature value falling into class 4, a distance from faults falling into class 4, and so on. Such a method, although increasing considerably the number of computational nodes, is capable of providing an efficient and objective approach.

With regard to the number of hidden layers, a single layer is commonly used in landslide analysis applications. The number of nodes for this layer can be defined on the basis of the number of the input nodes and training cells by using empirical criteria (Yesilnacar & Topal 2005). The application of such criteria provided us with a wide range of values, going from 7 to 5,440 nodes, for subset 1. We thus choose two different configurations with 74 (number of input nodes increased by 1) and 140 (about double) nodes in the hidden layer, respectively. Lastly, one node is used in the output layer, corresponding to the output value of susceptibility at each cell; output values range from 0, for minimum susceptibility, to 1, for maximum susceptibility, and ultimately represent the probability of slope failure for the cell. The overall network structures are denoted as $73 \times 74 \times 1$(RN1) and $73 \times 140 \times 1$(RN2) and are shown in Figure 3(b).

Table 1 | Runs carried out in the analysis. Configurations differ for subset used in the training phase, number of nodes in the hidden layer, and used back-propagation algorithm

		Dataset description								
Run	**Training dataset**	**Landslide cells**	**%**	**No landslide cells**	**%**	**TOT cells**	**%**	**Network**	**No nodes hidden layer**	**Back-propagation algorithm**
NN1	Subset 1	399	43	1,741	20.1	2,140	22.3	RN1	74	SCG
NN2	Subset 1	399	43	1,741	20.1	2,140	22.3	RN2	140	SCG
NN3	Subset 2	464	50	464	5.4	928	9.7	RN2	140	SCG
NN4	Subset 3	464	50	928	10.7	1,392	14.5	RN2	140	SCG
NN5	Subset 3	464	50	928	10.7	1,392	14.5	RN2	140	GDM

Table 2 | Processing of landslide inducing factors and landslide variable in classes

ID	Variable	No. classes
1	Aspect	9
2	Curvature	5
3	Distance from faults	10
4	Distance from stream network	10
5	Lithology	6
6	Hydrological parameter *n*	5
7	Pedology	8
8	Slope	6
9	Land use	14
	Total number of classes	73

(3) *Network training*

Among all the back-propagation algorithms available in the literature, two of the most suitable to treat a large amount of data are the GDM (gradient descend with momentum) and the SCG (scaled conjugate gradient) algorithms. Both the algorithms were used in the simulations according to the configurations described in Table 1, in order to compare the performances of both the models. The chosen activation function is a sigmoid function (sgm), which returns values ranging from 0 to 1.

(4) *Classification*

In order to estimate the landslide susceptibility, all the landslide-inducing factors are fed into the designed MLP network. The network returns the susceptibility values at each cell grid on the basis of the weights found during the training phase. For each cell, the relative position in the grid structure is recorded and used to reconstruct the susceptibility grid.

All the used configurations are summarized in Table 1, for a total of five runs indicated as NN followed by an ordinal number (first column). Each run uses a different combination of training dataset (subset 1, subset 2, and subset 3) (second column), network structure (RN1 and RN2) (fourth column) and back-propagation algorithm (GDM and SCG) (last column). The number of landslide and no landslide cells used during the network training is reported in the third column while the fifth column reports the number of nodes in the hidden layer.

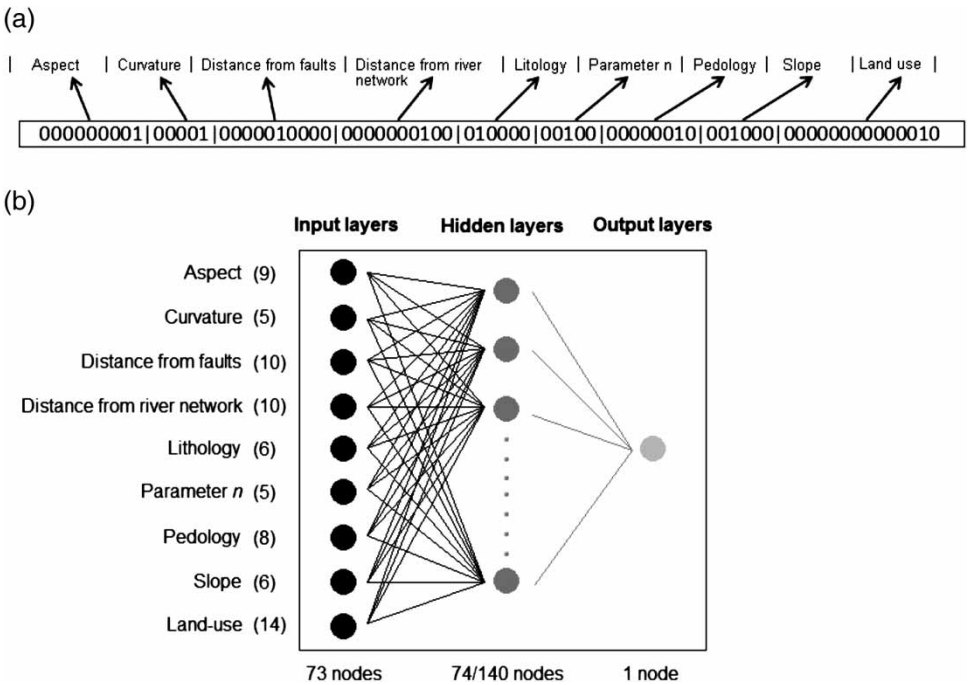

Figure 3 | (a) Representation of the binary string used to generate the input vector (modified from Ermini *et al*. (2005)) and (b) the overall network structures used for the simulations.

RESULTS AND DISCUSSION

In this section, the performances of the different experiments are compared and discussed in order to analyze the effects of network design, training subsets, and back-propagation algorithms on the results. The quantitative evaluation of goodness of fit of the models and their comparison are provided by the ROC curves and their relative AUC, all shown in Figure 4. As previously stated, the higher the AUC, the better the model is in terms of reproducing the occurrences. In particular, Hosmer & Lemeshow (2000) suggested the following general rule in order to interpret the capability of the model: $AUC = 0.5$ suggests no discrimination; $0.7 \leq AUC < 0.8$ is considered acceptable discrimination; $0.8 \leq AUC < 0.9$ is considered excellent discrimination; $AUC \geq 0.9$ is considered outstanding discrimination.

Effects of network design

Runs NN1 and NN2 use the same subset for training phase (subset 1), the same back-propagation algorithm (SCG) but different network configurations, RN1 and RN2, respectively, which are characterized by 74 and 140 nodes in the respective hidden layers. Both the networks provide acceptable results in terms of fit of model, with AUC values equal to 0.707 and 0.727, respectively, both greater than 0.7 indicated as the lower limit for acceptable discriminations (Hosmer & Lemeshow 2000). Even if RN2 provides a more efficient model (increase by 4% of AUC, Figure 4), it is a more complex network, with a greater number of nodes which leads to higher computational costs.

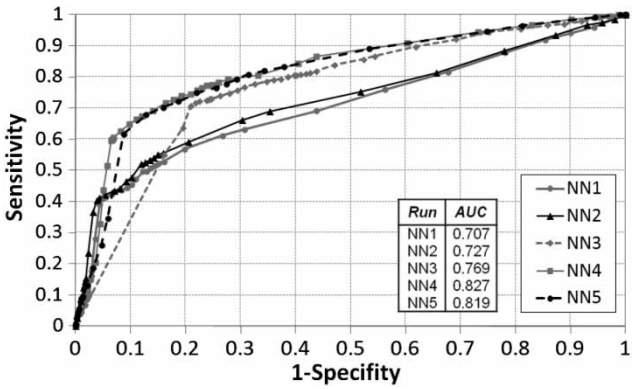

Figure 4 | ROC curves obtained for the five configurations of ANNs and the corresponding values of area under the curve (AUC).

Effects of the training subset

Runs NN2 and NN3 differ in using two different subsets for the training phase, methodological (subset 1) and random (subset 2), respectively. As stated, the methodological criterion proposed ensured the training pixels were selected over the entire basin, without fixing the number of landslide and no landslide cells. The resulting percentages of selected pixels were 43% among the landslide cells and 20.1% among the no landslide cells, for a total of 2,140 cells, corresponding to 22.3% of the entire basin area. On the other hand, the random criterion selected the cells by fixing an equal number for both landslide and no landslide cells (ratio 1:1), resulting in 50 and 5.4% of the respective areas, randomly distributed over the basin. In this case, a total of 928 cells were selected, corresponding to 9.7% of the entire basin area.

By comparing the results of simulations (Figure 4), it is interesting to observe that run NN3, which uses the random criterion of data selection, provides a value of AUC higher than run NN2 (0.769 versus 0.727), an improvement of 8.2% for the model performance. In other words, a smaller sample (9.7% versus 22.3% of the entire dataset), but more representative of the modeled phenomenon (50% of landslide occurrences versus 43%), allows the network to better learn the data structure and thus better reproduce the spatial patterns.

In subset 3, used for run NN4, the number of landslide cells (464) does not change, but the ratio of landslide/no landslide cells is equal to 1:2 and the number of no landslide cells is doubled to 928 (10.7%). A total of 1,392 elements are selected, corresponding to 14.7% of the entire basin area. NN4 results show a further improvement in the AUC value (equal to 0.827 with an increase of 11.6% in model performance with respect to NN3), suggesting that a correct choice of ratio between landslide and no landslides elements (in this case 1:2) is also important.

Effects of the back-propagation algorithm

Runs NN4 and NN5 use the same subset for the training phase (subset 3), but different back-propagation algorithms,

SCG and GDM, respectively. Both algorithms are the most used in such analyses because they are suitable in managing large datasets. The results shown in Figure 4 indicate that the two algorithms provide almost the same performances, with an AUC of NN5 slightly lower than NN4, equal respectively to 0.819 and 0.827.

Although the AUCs are very similar, the two runs have very different computational performances. The GDM algorithm required a much longer training phase, with a number of iterations (epochs) equal to ten times those required in SCG, 3,000 versus 300. Other differences are found in the spatial patterns of the output values, discussed in the following section.

Susceptibility maps

In order to obtain the final susceptibility maps and make their comparison easier, the slope failure probability distribution obtained by each ANN application was classified into five levels of probability. Classes are defined as follows: very low (0–0.1), low (0.1–0.25), medium (0.25–0.45), high (0.45–0.7), very high (0.7–1) susceptibility. The choice of these class breakpoints is based on the literature, and considers values greater than 0.4–0.5 as highly susceptible (Lee *et al.* 2003; Ermini *et al.* 2005; Melchiorre *et al.* 2008).

Figure 5 shows the relative frequency distribution of the susceptibility areas over the five classes and for the five configurations. The susceptibility values returned by the models are mostly distributed over two of the five available classes, i.e., the basin is mostly classified either as very low susceptibility or very high susceptibility to landsliding. In particular, for runs NN1 and NN2, almost 80% of the basin area falls in the lowest class and about 10% in the highest class. The remaining 10% is equally distributed among the other three classes. The corresponding maps, together with the locations of historical events (depicted as polygons), are shown in Figure 6 (NN1 and NN2); the two maps are very similar. Although the percentage of high susceptibility area is close to the percentage of landslide area (about 10%), there is not a perfect agreement between the two areas, as measured in the ROC curves. Most of the basin area is classified as very low, the remaining classes appeared to be spread over the basin as 'small spot' areas, mostly corresponding to single cells. A visual comparison between the two maps suggests that increasing the number of nodes in the hidden layer from 74 (NN1) to 140 (NN2) does not significantly change the reproduced pattern.

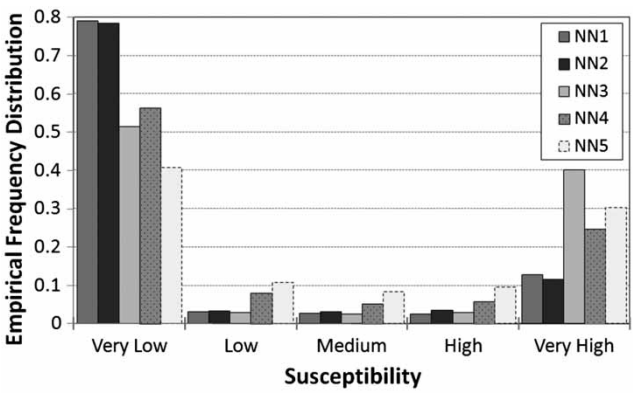

Figure 5 | Empirical frequency distribution of susceptibility values over five classes of susceptibility for the five configurations of ANNs.

Run NN3 classifies 40% of the basin area in the very high class and 50% in the very low class; again, the remaining 10% is equally distributed among the other three classes (Figure 5). Visually there is a good agreement between the very high susceptibility areas and the locations of historical events (polygons) (Figure 6, map NN3); however, the areas falling in the highest class are too large compared to the landslide polygons, generating a large number of false positives. This particular result is due to the choice of subset 2 in the training phase, where the network learned more about the landslide areas than the no landslide areas (in subset 2, 50% of landslide cells versus 5.4% of no landslide cells were used, see Table 1).

Run NN4, which uses subset 3 for the training phase and thus a greater number of no landslide cells, provides a different classification. The relative frequency distribution depicted in Figure 5 shows that in NN4, 56% of the basin is classified as very low susceptibility, almost 25% as very highly susceptibility and then 8% as low, 5% as medium, and 6% as high susceptibility. Also, the corresponding map presented in Figure 6 shows a spatial distribution of the very high susceptibility areas over the basin different from the previous maps, in very good agreement with the landslide locations.

Finally, in run NN5, which differs from the last case only in the back-propagation algorithm (GDM instead of SCG), the basin area is classified with values of susceptibility falling within all the five classes more homogeneously. In particular,

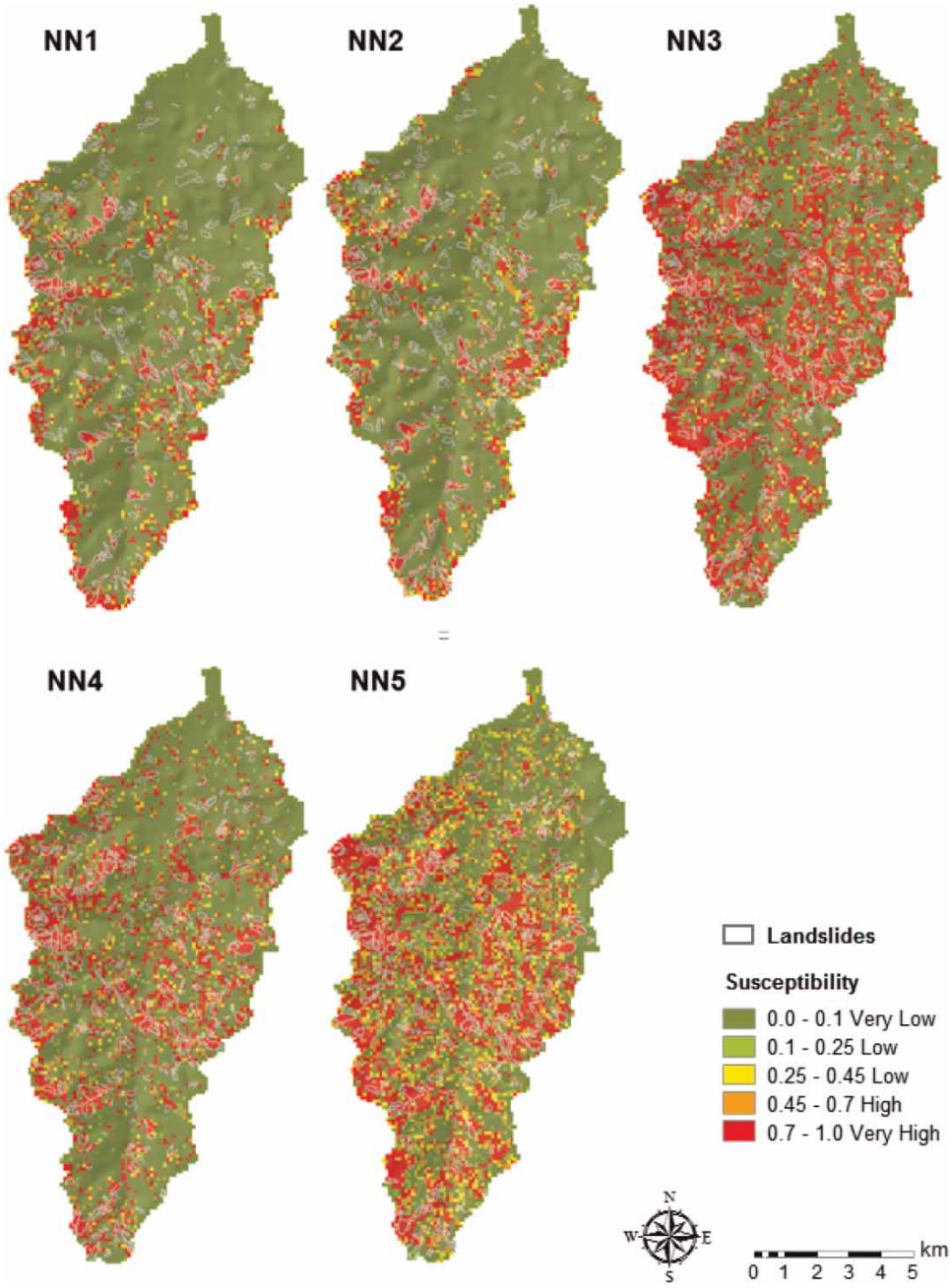

Figure 6 | Susceptibility maps of the five configurations of ANNs. Historical landslide locations are also reported in each map.

about 40% of the basin area falls within the very low class, 30% within the very high class, and then 11, 9, and 10% within the low, medium, and high classes (Figure 5). The corresponding map in Figure 6 shows again a very good agreement with landslide locations but a large number of false positive hits. In fact, the AUC value is slightly lower than the case described above.

Remarks

Overall, all configurations have shown acceptable results in terms of capability to correctly detect the occurrence or non-occurrence of predefined events; this is confirmed by the 'acceptable' or 'good' AUC values, which are within the range of values obtained by other studies (e.g., Yesilnacar &

Topal (2005)). Runs NN4 and NN5 seem to provide the best results, both in terms of fit of model and spatial patterns. These results confirm the importance of selecting a proper sample for the network training, in order to understand and reproduce the correct structure of data. In fact, the results demonstrated that a larger sample does not necessarily lead to higher model performance, while subsets with higher percentages of landslide area do. Increasing the number of nodes in the hidden layer slightly improves the model performances, while the choice of the back-propagation algorithm has to be evaluated carefully to improve the spatial results, but does not increase the model capability.

CONCLUSIONS

The use of ANNs in landslide susceptibility analyses has become a valid alternative to the use of statistical methods, due to the particular capability in analyzing spatial correlation data, without requiring strict statistical assumptions.

The application of ANNs to a given area requires the definition of several model strategies that need to be carefully set with regard to the definition of the model structure and the selection of dataset for a proper learning phase. This study has provided an investigation of model strategies by applying different MLP network configurations to a small Sicilian basin, in order to evaluate the effects of each strategy on the resulting landslide susceptibility maps. In particular, the following issues were analyzed: data selection for the network training, number of nodes in the hidden layer, back-propagation algorithms. Quantitative evaluation of model performances and their comparison were made by means of the AUC.

On the basis of the AUC values, according to the general rule suggested by Hosmer & Lemeshow (2000), all network configurations showed an acceptable capability to correctly discriminate the data experiencing landslides from those not experiencing landslide.

The higher number of nodes in the hidden layer leads to a slight improvement of model performances; however, this may not justify the increase of the computational costs given by a more complex network.

With regard to the strategies in selecting the data for the network training, it was demonstrated that the largest sample does not lead to the most performing model, as claimed by Lee *et al*. (2003); however, in contrast to that same work, it was demonstrated that the percentage of selected landslide area, related to the entire basin area (and not only landslide area), plays a more significant role. Moreover, it was confirmed that the commonly used random criterion of data selection (Ermini *et al*. 2005; Lee & Pradhan 2007; Caniani *et al*. 2008) leads to a more performing learning phase when compared with a methodological approach.

Finally, the use of two of the most known back-propagation algorithms, GDM and SCG, did not lead to significant differences in model performances. However, they lead to a different distribution of susceptibility values within the range of validity, with the GDM algorithm providing output results within all the susceptibility classes at considerable percentages.

Moreover, the results showed that MLP network models are capable of providing satisfactory agreement with the existing landslide location data, which have been classified within the higher susceptibility classes; this occurs especially in those cases that use the random criterion for data selection.

Although the overall results were satisfactory, it is worth pointing out that ANN models do not offer the possibility of making explicit direct considerations on the role of each landslide-inducing factor. In fact, it is difficult to follow the internal learning process (Lee *et al*. 2003) and to understand the physical relationship between factors and modeled phenomenon, which is instead possible to define in the statistical methods.

ACKNOWLEDGEMENTS

The work has been funded by Regional Sicilian Government (Italy), 'Linea di intervento 4.1.1.1 – POR FESR Sicilia 2007–2013' under Project SESAMO – SistEma informativo integrato per l'acquisizione, geStione e condivisione di dati AMbientali per il supportO alle decisioni.

REFERENCES

APAT 2005 *La realizzazione in Italia del Progetto Corine Land Cover2000*. APAT, Rapporti 36/2005, Rome, Italy.

Arnone, E., Noto, L. & Francipane, A. 2012 Landslide susceptibility mapping: a comparison of logistic regression and neural networks methods in a small Sicilian catchment. *10th International Conference on Hydroinformatics*, HIC 2012, Hamburg, Germany.

Arnone, E., Noto, L. V., Lepore, C. & Bras, R. L. 2011 Physically-based and distributed approach to analyze rainfall-triggered landslides at watershed scale. *Geomorphology* **133**, 121–131.

Bai, S.-B., Wang, J., Lu, G.-N., Zhou, P.-G., Hou, S.-S. & Xu, S.-N. 2009 GIS-based and data-driven bivariate landslide-susceptibility mapping in the three Gorges Area, China. *Pedosphere* **19**, 14–20.

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., Cardinali, M., Detti, R., Guzzetti, F., Pasqui, V. & Reichenbach, P. 1991 GIS techniques and statistical models in evaluating landslide hazard. *Earth Surf. Proc. Land.* **16**, 427–445.

Caruana, R., Lawrence, S. & Giles, L. 2001 Overfitting in neural nets: backpropagation, conjugate gradient, and early stopping. In: *Advances in Neural Information Processing Systems 13* (T. K. Leen, T. G. Dietterich & V. Tresp, eds). The MIT Press, Cambridge, MA, pp. 402–408.

Chung, C. F., Fabbri, A. G. & Van Westen, C. J. 1995 *Multivariate Regression Analysis for Landslide Hazard Zonation. Geographical Information Systems in Assessing Natural Hazards*. Kluwer, The Netherlands, pp. 107–133.

Di Piazza, A., Conti, F. L., Noto, L. V., Viola, F. & La Loggia, G. 2011 Comparative analysis of different techniques for spatial interpolation of rainfall data to create a serially complete monthly time series of precipitation for Sicily, Italy. *Int. J. Appl. Earth Obs.* **13**, 396–408.

Doglioni, A., Fiorillo, F., Guadagno, F. & Simeone, V. 2012 Evolutionary polynomial regression to alert rainfall-triggered landslide reactivation. *Landslides* **9**, 53–62.

Ermini, L., Catani, F. & Casagli, N. 2005 Artificial Neural Networks applied to landslide susceptibility assessment. *Geomorphology* **66**, 327–343.

Falaschi, F., Giacomelli, F., Federici, P. R., 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.

Fawcett, T. 2006 An introduction to ROC analysis. *Pattern Recogn. Lett.* **27**, 861–874.

Fierotti, G. & Ballatore, G. P. 1988 Carta dei suoli della Sicilia. scala 1:250.000, Universita degli Studi di Palermo, Unione delle Camere di commercio, industria, artigianato e agricoltura della Regione Sicilia.

Giustolisi, O. & Savic, D. A. 2006 A symbolic data-driven technique based on evolutionary polynomial regression. *J. Hydroinform.* **8**, 207–222.

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.

Hansen, A. 1984 Landslide hazard analysis. In: *Slope Instability* (D. Brunsden & D. B. Prior, eds). Wiley & Sons, New York, pp. 523–602.

Haykin, S. 1994 *Neural Networks: A Comprehensive Foundation*. Prentice Hall, Englewood Cliffs, NJ.

Hines, J. W. 1997 *Fuzzy and Neural Approaches in Engineering*. John Wiley & Sons, New York, 210 pp.

Hosmer, D. W. & Lemeshow, S. 2000 *Applied Logistic Regression*. John Wiley & Sons, New York.

Lee, S. & Pradhan, B. 2007 Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. *Landslides* **4**, 33–41.

Lee, S., Ryu, J. H., Min, K. & Won, J. S. 2003 Landslide susceptibility analysis using GIS and artificial neural network. *Earth Surf. Proc. Land.* **28**, 1361–1376.

Lee, S., Choi, J. & Woo, I. 2004a The effect of spatial resolution on the accuracy of landslide susceptibility mapping: a case study in Boun, Korea. *Geosci. J.* **8**, 51–60.

Lee, S., Ryu, J. H., Won, J. S. & Park, H. J. 2004b Determination and application of the weights for landslide susceptibility mapping using an artificial neural network. *Eng. Geol.* **71**, 289–302.

Lee, M., Wang, S. & Lin, T. 2010 The effect of spatial resolution on landslide mapping – a case study in Chi-Shan River Basin, Taiwan. *31st Asian Conference on Remote Sensing*, Hanoi, Vietnam.

Lepore, C., Kamal, S. A., Shanahan, P. & Bras, R. L. 2012 Rainfall-induced landslide susceptibility zonation of Puerto Rico. *Environ. Earth Sci.* **66**, 1667–1681.

Lippmann, R. P. 1987 Introduction to computing with Neural Nets. *IEEE ASSP Mag.* **4**, 4–22.

Lo Conti, F., Noto, L. V., Cannarozzo, M. & La Loggia, G. 2007 Regional frequency analysis of extreme precipitation in Sicily, Italy. *2nd International Workshop on Hydrological Extremes: Variability in Space and Time of Extreme Rainfalls*, Floods and Droughts, Cosenza.

Melchiorre, C., Matteucci, M., Azzoni, A. & Zanchi, A. 2008 Artificial neural networks and cluster analysis in landslide susceptibility zonation. *Geomorphology* **94**, 379–400.

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

P.A.I. 2006 Piano Stralcio di Bacino per l'Assetto Idrogeologico (P.A.I). Bacino Idrografico del Torrente Timeto (012), Assessorato Territorio e Ambiente, Regione Sicilia.

Pradhan, B. & Lee, S. 2009 Landslide risk analysis using artificial neural network model focussing on different training sites. *Int. J. Phys. Sci.* **4**, 01–015.

Tian, Y., Xiao, C., Liu, Y. & Wu, L. 2008 Effects of raster resolution on landslide susceptibility mapping: a case study of Shenzhen. *Sci. China Ser. E* **51**, 188–198.

Varnes, D. & IAEG 1984 *Landslide Hazard Zonation: A Review of Principles and Practice*. UNESCO Press, Paris, p. 63.

Viola, F., Noto, L. V., Cannarozzo, M. & La Loggia, G. 2011 Regional flow duration curves for ungauged sites in Sicily. *Hydrol. Earth Syst. Sci.* **15**, 323–331.

Werbos, P. J. 1994 *The Roots of Backpropagation: From Ordered Derivatives to Neural Networks and Political Forecasting*. John Wiley & Sons, New York.

Wilson, J. P. & Gallant, J. C. 2000 *Terrain Analysis: Principles and Applications*. John Wiley & Sons, New York.

Yesilnacar, E. & Topal, T. 2005 Landslide susceptibility mapping: a comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). *Eng. Geol.* **79**, 251–266.

Yilmaz, I. 2010 The effect of the sampling strategies on the landslide susceptibility mapping by conditional probability and artificial neural networks. *Environ. Earth Sci.* **60**, 505–519.

First received 30 October 2012; accepted in revised form 20 March 2013. Available online 19 April 2013