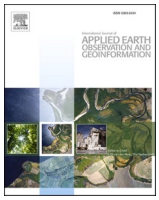
International Journal of Applied Earth Observations and Geoinformation 104 (2021) 102549

Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

journal homepage: www.elsevier.com/locate/jag



Data–driven classification of landslide types at a national scale by using Artificial Neural Networks

Gabriele Amato, Lorenzo Palombi *, Valentina Raimondi

"Nello Carrara" Applied Physics Institute of the National Research Council of Italy (CNR-IFAC), Sesto Fiorentino, Italy

ARTICLE INFO

Keywords:
Data-driven classification
Artificial Neural Network
Machine Learning
Landslide inventory
Landslide type
Geospatial modelling

ABSTRACT

Classification of landslide type is an essential step in risk management, although is often missing in large inventories. Here we propose a novel data-driven method that uses easily accessible morphometric and geospatial input parameters to classify landslides type at a national scale in Italy by means of a shallow Artificial Neural Network. We achieved an overall True Positive Rate of 0.76 for a five-class overall classification of over 275,000 landslides as (1) rockfall/toppling, (2) translational/rotational slide, (3) earth flow, (4) debris flow, and (5) complex landslide. In general, the model performance is very good in the entire national territory, with large areas reaching F-score higher than 0.9. The method can be applied to any polygonal inventory, as those produced by automatic mapping procedures from Earth Observation imagery, in order to automatically identify the types of landslides.

1. Introduction

The first step in landslide hazard assessment is to map the exact location of gravitational processes and to define the type of movement (Mantovani et al., 1996). In engineering geology, classifying landslides by type allows defining of the most suitable instrumental investigations needed to study in depth the phenomenon (Popescu, 2001).

At regional scale, clusters of landslides of the same type may be associated with similar geo-environmental features (Rosi et al., 2018) which can be, for instance, indirectly inferred by observing the type of landslide. In addition, different types of failures may evolve in a different way under the effect of global warming (Gariano and Guzzetti, 2016). Furthermore, different threats can arise depending on the type of mass movement thus, its classification is an essential step for planning adequate countermeasures at site and regional scales (Guzzetti et al., 1999).

The failure typology is usually defined in accordance with a classification system commonly accepted in literature, such as: the 1978 "Varnes Classification System" (Varnes, 1978) and relevant updates (Dikau et al., 1996; Hungr et al., 2014; Cruden and Varnes, 1996) ; those proposed by Hungr et al. (2001) and Hutchinson (1988); or some simplified versions of the aforementioned systems as proposed by Guzzetti et al. (2012). In addition, detailed classifications have been proposed for specific types of landslide (e.g. Li and Mo, 2019).

During the production of landslide inventories, mass movement type is classified by means of field survey or image interpretation done by an expert, with big effort in terms of time and human resources (Prakash, et al. 2020). In the last decades, the increased availability of satellite products has fostered the production of an increasing number of landslide inventories by means of Automatic or semi-automatic Mapping (AM) methods, reducing the time needed for the production of a landslide map (Guzzetti et al., 2012). Nevertheless, AM approaches usually suffer from some difficulties in automatically defining the type of failure with the same proficiency they detect them (Barlow et al., 2006). As a result, landslide inventories are often inaccurate for one or more of the following reasons: (i) they ignore the coexistence of more landslide types in a given area (Moosavi et al., 2014); (ii) they *a priori* refer only to a single – often generic – landslide type (e.g. "shallow landslides", Amato et al., 2019); (iii) they perform a simplified classification (e.g. "shallow vs. deep") (Mezaal and Pradhan, 2018). Only few studies demonstrated the possibility to determine landslide types automatically and according to a complex classification system (Barlow et al., 2006; Martha et al., 2010, 2011; Taalab et al., 2018). These have in common one or more of the following aspects: (i) landslides type classification is usually combined with landslides mapping from remote sensing imagery; (ii) rules set by the user are applied to drive the failure type classification; iii) they

* Corresponding author at: "Nello Carrara" Institute of Applied Physics - National Research Council of Italy (CNR-IFAC), Via Madonna del Piano, 10, I-50019 Sesto Fiorentino, Firenze, Italy.
E-mail address: l.palombi@ifac.cnr.it (L. Palombi).

https://doi.org/10.1016/j.jag.2021.102549
Received 4 May 2021; Received in revised form 16 July 2021; Accepted 14 September 2021
Available online 22 September 2021

adopt a combination of spectral, terrain and morphometric parameters, and (iv) they usually deal with small inventories (<100 landslides). Only Taalab et al. (2018) adopts a large inventory (>30,000) to produce a landslide susceptibility map for multiple types of landslides.

In short, although the classification of mass movement types was demonstrated to be an important metric for risk management (Wood et al., 2020), such information is often missing in landslide inventories or is performed by means of supervised methods or limited to small inventories.

In this paper, we propose a novel, simple method for landslide type classification, which uses easily accessible input parameters and does not require the intervention of the user, and we apply it to a national-scale dataset. It relies on the use of shallow Artificial Neural Networks (ANNs). ANNs model the relationship between independent and dependent variables in a complex non-linear way and are intrinsically overparameterized. With respect to traditional methods, ANNs can model complex relations and thanks to overparameterization, they are not very sensitive to collinearity problems (De Veaux and Ungar, 1994). This typically ensures a greater robustness of the predictive performances (Garg and Tai, 2012) and the possibility of considering a greater number of predictive variables. On the other hand, due to over-parameterization, they are not typically used for a model interpretation but mainly for predictive purposes. The method contains several novel elements. First, it relies on the exploitation of the information extracted from a Digital Elevation Model (DEM) and from the inventory (e.g. landslide shape-related indices), instead of using aerial images and spectral information. Indices like compactness and length-to-width ratio have been included in the analysis, posing this work in continuity with past experiments concerning the link between landslide type and its shape (Taylor et al., 2018). Secondly, the analysis includes the spatial distribution of the landslides, proposing an original approach rarely discussed in the literature. Thirdly, it relies on an existing inventory to extract the information instead of generating an *ad hoc* landslide map. Finally, the method does not require any pre-set rule to drive the classification since the ANN identifies the most important parameters to optimise the classification performances.

In conclusion, to the best of authors' knowledge, this is the first paper proposing a simple, data-driven model of landslide type classification since it does not require any pre-set rule to drive the classification and does use only the information extracted from a DEM and from landslide shape-related indices extracted from traditional inventories. The method can also be tested in other geographical areas where a DEM and a polygonal landslide inventory are available. It can also find potential use in the revision of classification accuracy of other inventories or to implement landslide type classification in those inventories in which this information is missing.

2. Material and methods

The workflow of the proposed method has been summarized in the flow chart shown in Fig. 1. More detailed explanations of the methods used are provided in the following sections.

2.1. Dataset

The dataset we used to train the ANN and perform the classification of the landslide types was extracted from the IFFI Italian Landslide Inventory (Trigila et al., 2010). The IFFI catalogue counts over 620,000 mass movements, collected and validated by the means of aerial photo interpretation, historical documents and field surveys, while their type classification refers mainly to Varnes (1978) and Cruden and Varnes (1996) systems. From the full IFFI catalogue, we selected landslides with associated polygonal shape data among the ones classified as: (1) rockfall/toppling, (2) translational/rotational slide, (3) earth flow, (4) debris flow, and (5) complex landslide, which were the most populated landslide categories in the inventory and are among the most common types of landslide. These five mutually exclusive classes can be reasonably considered also as jointly exhaustive of the entire dataset since the total number of landslides in classes other than these represents only 0.07% of the total.

2.2. Model parameters

The total number of parameters considered by the network in order to perform the landslide classification is 53. They were associated with each polygon of the landslide dataset. These parameters can be divided into three groups: (1) Spatial parameters, providing information on the geographical position of each failure and on the other landslides in its vicinity; (2) Terrain parameters, constituted by some of the classical DEM-derived indices largely used in landslide geospatial models; (3) Shape-related parameters, which represent the geometric properties of the landslide polygon shape.

2.2.1. Spatial parameters

The occurrence of a landslide in a certain position implies the presence of peculiar geomorphological and geological conditions, co-existing at the failure point (Rosi et al., 2018; Guzzetti, 2000). Therefore, as a first approximation, the combination of these conditions can be resumed by providing its geographic location. Furthermore, we assumed that the occurrence of a type of landslide in a given area may imply the presence of other, same-type, failures in its vicinity, as in the cases of a multiple-occurrence landslide event (Crozier, 2005). In order to test this assumption, we computed for each landslide and fed to the ANN the North and East coordinates of the polygon centroid and a series of

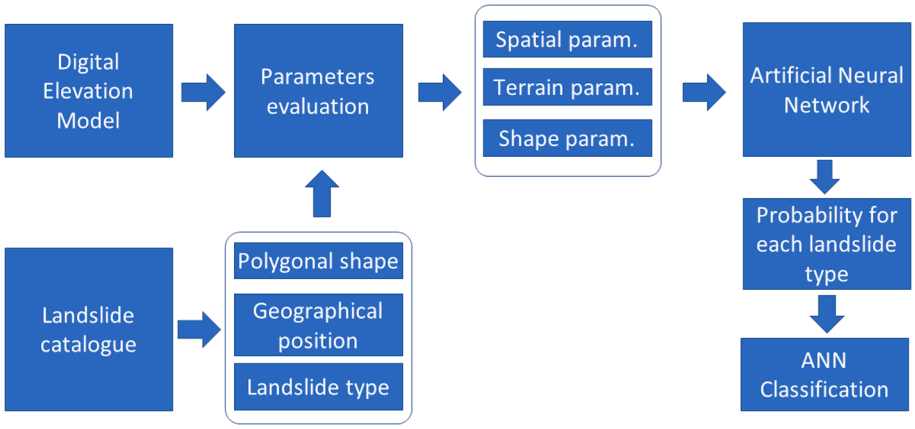


Fig. 1. Workflow of the proposed method.

parameters, described below, which aimed at providing information on the types of the landslides located in the vicinity of each failure. These parameters refer to landslides located within the given radius from a landslide and they are:

"NearSlopeN": total number of landslides.

"NearMeanDist": mean distance of the landslides.

"NearPerc*i*": ratio between the number of class *i* landslides (with *i* from 1 to 5) and the total number of landslides.

"NearNum*i*": number of class *i* landslides.

"NearSup*i*": the total area covered by class *i* landslides.

2.2.2. Terrain parameters

Here we used a 20 m DEM released by the Italian Institute for Environmental Research (ISPRA) in 2013. For each landslide polygon, we calculated the mean value and the standard deviation of the following DEM derivatives:

Aspect (Dai and Lee 2002), Elevation (Ayalew et al., 2004), General Curvature, Longitudinal Curvature and Tangential Curvature (Evans 1980), Planar and Profile Curvatures (Heerdegen and Beran, 1982), Slope (Zevenbergen and Thorne, 1987), Topographic Positioning Index (TPI, Pourghasemi et al., 2018), Topographic Roughness Index (TRI, Riley et al., 1999), Topographic Wetness Index (TWI, Beven and Kirkby, 1979).

2.2.3. Shape-related parameters

Here we included Area (A), Perimeter (P), and a list of geometrical indices computed in order to quantify the 2D and 3D characteristics of the landslide shape. These are:

"Compactness IPQ" (Osserman, 1978), "Compactness Moment of Inertia" (Bachi, 1999), "HullArea" (area of the convex hull), "HullPerimeter" (perimeter of the convex hull), "HullAreaRatio" (ratio between landslide area and convex hull area), "HullPerimRatio" (ratio between convex hull perimeter and landslide perimeter).

Furthermore, some parameters have been originally computed and considered for the first time in this work. These are based on two geometrical features, specifically "Direction of maximum slope" and "Centroidal Principal moments and axes".

The first is the direction of the line that links the highest part of the landslide to the lowest one and represents an approximation of its expected direction of movement. We calculated "the highest part" as the centroid of the contour points whose elevation is greater than the 70% of the heights range, and "the lowest part" as the centroid of the ones whose elevation is lower than the 30% of the heights range.

"Centroidal Principal moments and axes" were calculated on the basis of the landslide area and shape. The major/minor principal moment is the maximum/minimum moment of inertia among those evaluated by considering all possible axes passing through the centroid of the area. The major/minor principal axis of the shape corresponds to the major/minor principal moment. The minor principal axis corresponds to the main linear orientation of the object, i.e. to the direction along which the shape is more elongated, and it is perpendicular to the major one.

In detail, the original shape-related parameters are:

"SlopeLengthCos": cosine of the angle between the direction of maximum slope and the minor principal axis. The parameter aims at evaluating whether the direction of movement of a landslide is close to the direction along which the shape is more elongated. This parameter was considered because several contiguous small landslides might be sometimes mapped as a single, larger polygon, resulting elongated perpendicularly to its direction of movement. This parameter, instead, allows to distinguish between a narrow, elongated landslide and a large, short-running one.

"OblongityMoments": it was calculated as the square root of the ratio between the major and minor principal moments. This parameter aims at quantifying the difference between the shape distribution along the main orientation and its perpendicular.

"SqrtMomentsMaxSlope": it was calculated as the square root of the ratio between the centroidal moment wrt the perpendicular to the maximum slope direction and the centroidal moment wrt the maximum slope direction. This parameter aims at quantifying the difference between the shape distribution along the direction of the landslide's maximum slope and its perpendicular.

"Number of holes within a landslide polygon": this factor aims at reproducing the presence, within a large mass movement, of smaller failures mapped separately generating "holes". Intuitively, this circumstance can more easily occur in complex landslides (class 5), slow-moving flows (class 3) and slides (class 2) than in rockfalls (class 1) or debris flows (class 4).

A panoramic view and a short description of all the 53 parameters is provided in Table 1.

2.3. Modelling procedure

We adopted this procedure: (i) we selected a landslide dataset from the IFFI inventory, composed by a number N of samples (landslide polygons); we split them into training, validation and test subsets; (ii) we produced a list of a number M of parameters potentially useful to distinguish between different landslide types; (iii) we associated 53 values (one for each parameter) to each sample by creating a N by M matrix and used it as an input dataset for the ANN; to each sample, the corresponding class of failure type is also associated as expected output; (iv) we set the optimal ANN design and we trained, validated and tested the network over the respective subsets. The complete dataset counted 275,749 landslides, among which 8313 (3.02%) belong to class 1, 132,561 (48.07%) to class 2, 67,151 (24.35%) to class 3, 14,845 (5.38%) to class 4 and 52,879 (19.18%) to class 5.

ANNs training does not require an equal number of samples per class (balanced dataset). The machine learning methods for supervised classification, however, in the case of an imbalanced dataset tend to provide better performances for the more most populated classes to the detriment of with respect to the less least populated ones. In particular, there is a tendency to over-classify the samples in the most populated classes with the consequent misclassification of the samples belonging to the least populated classes. Even general performance metrics tend to be biased by the presence of class imbalance. In order to overcome these problems, it is possible to consider algorithm-level methods or simpler data-level methods (Johnson and Khoshgoftaar, 2019). The latter involve both random under-sampling of the most numerous classes and random over-sampling with repetition of the least numerous ones (Van Hulse et al., 2007).

Thus, we preliminarily balanced the number of samples per class as follows. If N is the total number of samples, the number Q of samples per class is calculated by dividing N by the number of classes. For each class, if the total number of available landslides is greater than Q, we randomly selected a number Q of them without repetition. If the total number of available landslides is smaller than Q, a number Q of them is randomly selected with repetition. In order to preserve their independence, the training, validation and test datasets are randomly extracted by the N samples dataset with a constraint of complementarity: the same landslide can belong to only one of the three datasets.

In the end, we set a fixed number of 32,000 samples per class, this number being chosen according to the criteria explained in Section 2.4, so that the whole dataset fed to the network counted 160,000 samples in total. This was split as follows: 70% of the samples were used to train the model, 15% for the validation and 15% for the test.

The training dataset was used by the network to optimise the weights and the bias of each node. During training, the ANN performed the same classification also on the validation dataset. By continuously comparing the results of these two datasets, model overfitting was avoided by using an early training stop criterium. The test dataset was an independent dataset used to test the reproducibility of the performances obtained on the first two datasets.

Table 1
Panoramic view and a short description of all the considered parameters.

Group	Name	Cod	Description
Terrain	Aspect	1	The average Aspect (Dai and Lee 2002) of the landslide polygon, calculated from the DEM at 20 m.
	General Curv.	2	The average General Curvature (Evans, 1980) of the landslide polygon, calculated from the DEM at 20 m.
	Longitudinal Curv.	3	The average Longitudinal Curvature (Evans, 1980) of the landslide polygon, calculated from the DEM at 20 m.
	Planar Curv.	4	The average Planar Curvature (Heerdegen and Beran, 1982) of the landslide polygon, calculated from the DEM at 20 m.
	Profile Curv.	5	The average Profile Curvature (Heerdegen and Beran, 1982) of the landslide polygon, calculated from the DEM at 20 m.
	Slope	6	The average Slope (Zevenbergen and Thorne, 1987) of the landslide polygon, calculated from the DEM at 20 m.
	Topographic Positioning Index (TPI)	7	The average TPI (Pourghasemi et al., 2018) of the landslide polygon, calculated from the DEM at 20 m.
	Topographic Roughness Index (TRI)	8	The average TRI (Riley et al., 1999) of the landslide polygon, calculated from the DEM at 20 m.
	Topographic Wetness Index (TWI)	9	The average TWI (Beven and Kirkby, 1979) of the landslide polygon, calculated from the DEM at 20 m.
	Tangential Curv.	10	The average Tangential Curvature (Evans, 1980) of the landslide polygon, calculated from the DEM at 20 m.
	Elevation	11	The average Elevation of the landslide polygon, calculated from the DEM at 20 m.
	Aspect STD	12	The standard deviation of the Exposure of the landslide polygon.
	Gen. Curv. STD	13	The standard deviation of the General Curvature of the landslide polygon.
	Long. Curv. STD	14	The standard deviation of the Longitudinal Curvature of the landslide polygon.
	Plan. Curv. STD	15	The standard deviation of the Plan Curvature of the landslide polygon.
	Prof. Curv. STD	16	The standard deviation of the Profile Curvature of the landslide polygon.
	Slope STD	17	The standard deviation of the Slope of the landslide polygon.
	TPI STD	18	The standard deviation of the TPI of the landslide polygon.
	TRI STD	19	The standard deviation of the TRI of the landslide polygon.
	TWI STD	20	The standard deviation of the TWI of the landslide polygon.
	Tan. Curv. STD	21	The standard deviation of the Tangential Curvature of the landslide polygon.
	Elevation STD	22	The standard deviation of the Elevation of the landslide polygon.
Shape	Slope-elongation cosine	23	Cosine of the angle between the direction of maximum slope and the minor principal axis. The parameter aims at evaluating whether the direction of movement (expressed as direction of maximum slope) of a landslide is close to the direction along which the shape is more elongated. This parameter was chosen because several contiguous small landslides might be sometimes mapped as a single, larger, polygon, resulting elongated perpendicularly to its maximum slope (or direction of movement). This parameter, instead, allows to distinguish between a narrow, elongated, landslide and a large, short-running, one.
	Area	24	Area of the landslide polygon.
	Perimeter	25	Perimeter of the landslide polygon.
	IPQ Compactness (Osserman, 1978)	26	Compactness of the landslide polygon shape measured as $4\pi A)/P^2$. A circular shape, the most compact, has an IPQ compactness equal to 1. Less compact shapes will values close to 0.
	Compactness (Bachi, 1999)	27	Compactness of the landslide polygon shape measured as $A^2/2\pi I_g$, where Ig is the is centroidal polar moment of the landslide shape. It is based on the ratio between the moment of inertia of a circle of the same area about its centre and the moment of inertia of the shape about its centroid (i.e. the centroidal polar moment). It varies between 0 and 1 (when the shape is a perfect circle).
	Oblongity	28	Square root of the ratio between the major and minor principal moments of the landslide polygon shape. By computing the ratio between the two principal moments, this parameter aims at quantifying the difference between the shape distribution along the main orientation (i.e. the minor principal moment) and its perpendicular (i.e. the major principal moment). This parameter is very similar to Oblongity, which was proposed by Bachi (1999) as a tool to study the noncompactness of a 2D object.
	SqrtMmtMax	29	It is calculated as the square root of the ratio between the centroidal moment with respect to the perpendicular to the maximum slope direction and the centroidal moment with respect to the maximum slope direction. This parameter aims at quantifying the difference between the shape distribution along the direction of the landslide's maximum slope and its perpendicular.
	Hull Area	30	Area of the convex hull of the landslide 2D shape.
	Hull Perimeter	31	Perimeter of the convex hull of the landslide 2D shape.
	Hull Area Ratio	32	Ratio between landslide area and convex hull area. It ranges between 0 and 1, where a value close to zero means that the landslide shape is particularly indented or concave.
	Hull Perimeter Ratio	33	Ratio between convex hull perimeter and landslide perimeter. It ranges between 0 and 1, where a value close to zero means that the

(continued on next page)

Table 1 (*continued*)

Group	Name	Cod	Description
			landslide shape is particularly indented or concave.
	N holes (number of holes within the landslide polygon)	34	This factor aims at reproducing the presence, within a large mass movement, of smaller failures mapped separately generating "holes" within the bigger polygon. Intuitively, this circumstance can more easily occur in complex landslides (type 5), slow-moving flows (type 3) and slides (type 2) than in rockfalls (type 1) or fast debris flows (type 4).
Spatial	North coordinate (Y)	35	North coordinate of the centroid of the landslide polygon.
	East coordinate (X)	36	East coordinate of the centroid of the landslide polygon.
	Near Landslides	37	Total number of landslides located within a given radius from the landslide polygon centroid.
	Mean Distance	38	Mean distance of the landslides located within a given radius from the landslide polygon centroid.
	% of class i ($i = 1…5$)	39–43	Ratio between the number of class i landslides and the total number of landslides within a circle of given radius from the landslide polygon centroid.
	N of class i ($i = 1…5$)	44–48	Number of class i landslides within a given radius from the landslide polygon centroid.
	Area of class i ($i = 1…5$)	49–53	The total area covered by class i landslides within a given radius from the landslide polygon centroid.

In order to build a statistically significant distribution of the model results and to calculate reliable mean and standard deviation of the performance metrics, we performed the landslide type classification training for 100 times. For each ANN training, the dataset composed of 160,000 was randomly generated from the 275,749 landslides of the complete dataset. ANN weights and biases are randomly initialised before each training also.

We represented the classification performance obtained by the network by means of the True Positive Rate (TPR) and the Receiver Operating Characteristic (ROC) curve, with related AUC (Area Under the Curve). Furthermore, we calculated False Positive Rate (FPR), False Negative Rate (FNR), True Negative Rate (TNR) and F-score (Tharwat, 2021).

2.4. ANN design

The ANN architecture is a two-layers feed-forward network optimised for multi-class classification. For the hidden layer, we considered a sigmoid activation function. The output layer is a "softmax layer", in which the outputs are normalised into probabilities proportional to the exponential of the input values. The network is trained by scaled conjugate gradient backpropagation. In order to limit any overfitting effect, we adopted an "early stopping by validation" training criterium. Cross entropy was used as loss function.

For each sample, the ANN assigns five probability values, namely one probability to belong to each landslide class. Each sample is thus classified as belonging to the class with the highest probability. The ANN hyperparameters have been tuned considering the number of available samples and trying to guarantee the best trade-off between performances and processing times.

2.5. Features importance analysis

In order to study the parameters importance for the classification task, we performed an a posteriori feature importance analysis. The method we adopted was the Permutation Feature Importance (PFI) (Putin et al., 2016). The PFI randomly permutates the values of one of the parameters within the ANN input matrix and keeps the others unchanged. The meaningful information of that parameter is thus made uncorrelated with the expected output, yet preserving the original distribution of parameter values. The more important is the parameter, the worse is the model performance as the permutation is applied. The PFI score for a parameter is here defined as the ratio between the TPRs without and with permutation.

Given the ANN architecture, we tuned the number of nodes in the hidden layer and the number of samples per landslide class. Since these two hyperparameters are highly related, they were tuned jointly: we calculated the ANN performances for all the possible combinations between 4000, 8000, 16000, 32000, 64,000 landslides per class and 1, 2, 4, 8, 12, 16, 20, 26, 32 nodes in the hidden layer, with a radius fixed at 1500 m for spatial parameters.

In Fig. 2 it is shown how the overall TPR varies for the three datasets and for the whole dataset (training + test + validation) as a function of the number of landslides per class and the number of hidden layers nodes.

As expected, TPR increases asymptotically with the number of the hidden layer nodes. This, however, occurs only if we consider a sufficiently large number of samples.

In order to avoid excessive oversampling, which is needed to balance the less populated classes, we operated a trade-off among the number of samples, the ANN performance and the computational burden. Considered that the dataset has 8313 landslides in the less populated class (class 1 - rockfall/toppling) and 132,561 in the most numerous one (class 2 - translational/rotational slide), we opted for setting 32,000 samples per class and 26 nodes for the ANN final design. This choice also permitted to maintain a moderate computing time. During the optimisation phase of the hyperparameters, we verified that oversampling did not provide any advantage or disadvantage to the performance of one class compared to those of the other classes. It should be noted that the number of landslides per class is an important parameter, but not critical for the applicability of the proposed method. Even when considering a low number of landslides, a minimum decrease in performance is observed. This is a good basis for the applicability of the proposed method to catalogues of landslides more limited than the considered one.

Another parameter to be tuned is the radius used to compute the spatial parameters. We calculated the ANN performance as the radius varied from 0 to 10 km. Fig. 3 shows that the mean of the overall TPR reaches its maximum for a 1500-m radius. This value was used in the optimised model.

3. Results

Once the hyperparameters were fixed, the ANN was trained for 100 times on randomized datasets, as described in section 2.3. Fig. 4 shows the TPR per class we obtained by applying the classification to the training, validation, test and whole datasets.

TPR values for training, test and validation datasets are very similar to each other. This demonstrates that training with early stopping by validation allows to minimise overfitting. The performance obtained on the whole dataset was taken as a representative of those obtained on the other three datasets since they are all very similar to each other. Table 2 shows the whole dataset mean confusion matrix (reported as the percentage of classified samples with respect to the total number of samples) and a list of performance indices achieved. All standard deviations of performance indices were less than 0.006.

We also evaluated the ROC curve of each class, which illustrates the

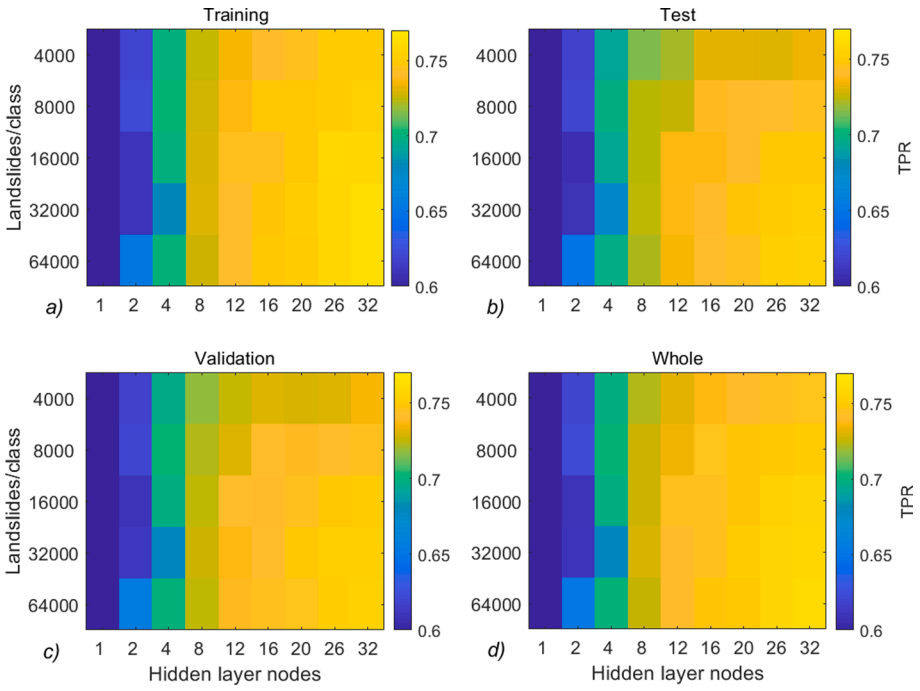


Fig. 2. Overall TPR as a function of the number of nodes in the hidden layer and of landslides per class: (a) training, (b) test, (c) validation, and (d) whole datasets.

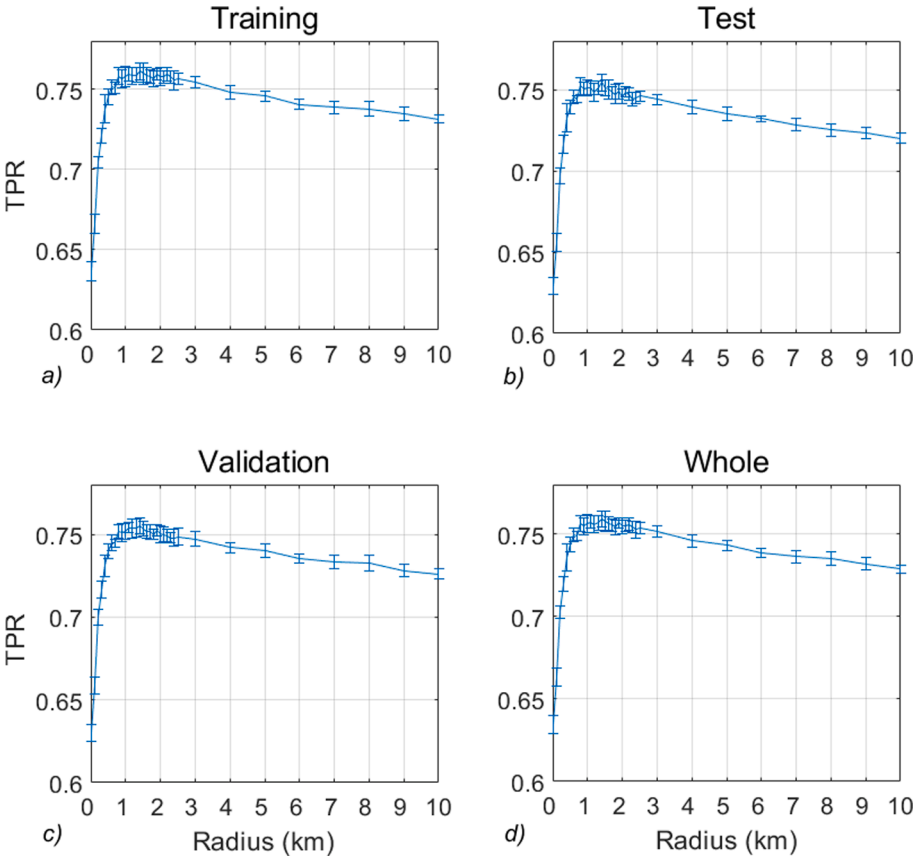


Fig. 3. Mean and standard deviation of the overall TPR as a function of the radius used to compute the spatial parameters: (a) training, (b) test, (c) validation and (d) whole datasets.

performances achieved in classifying the samples as belonging or not to that class for any possible classification probability threshold. Fig. 5 shows the mean ROC curve and AUC value for each class of the whole dataset. AUC standard deviations were less than 0.002.

In Fig. 6, we plotted, for each of the five classes the geographical distribution of the landslides and the corresponding mean value of the obtained F-scores. Both are evaluated on a regular 10-km by 10-km grid.

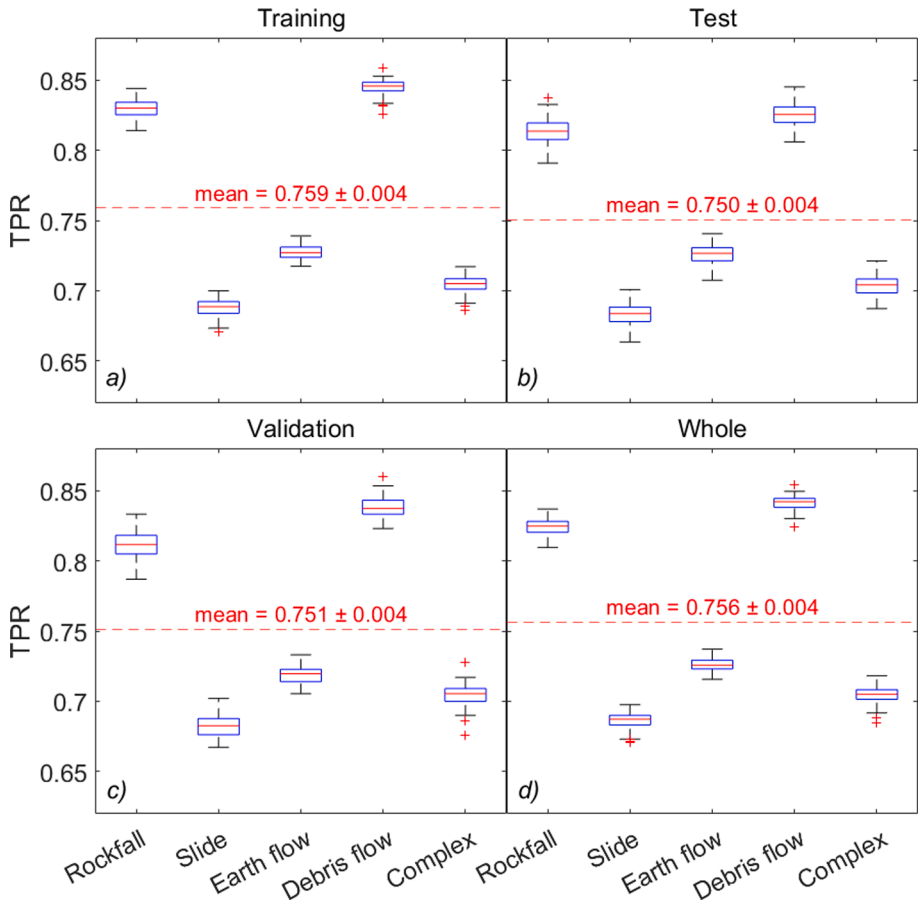


Fig. 4. Boxplots of the TPR per class: (a) training, (b) test, (c) validation and (d) whole datasets.

Table 2
Mean confusion matrix and performance indices for the whole dataset.

	PREDICTED						
Class	1	2	3	4	5	Sum	
1	16.49%	1.26%	0.61%	1.05%	0.59%	20%	
2	1.35%	13.72%	2.16%	0.72%	2.05%	20%	
3	0.47%	1.66%	14.51%	1.07%	2.29%	20%	**TRUE**
4	1.17%	0.54%	0.95%	16.82%	0.52%	20%	
5	0.98%	1.83%	2.19%	0.91%	14.09%	20%	
Sum	20.46%	19.01%	20.42%	20.57 %	19.54%	100%	
Metric	Per class					Overall	
TPR	0.824	0.686	0.726	0.841	0.704	**0.756**	
TNR	0.950	0.934	0.926	0.953	0.932	**0.939**	
FPR	0.050	0.066	0.074	0.047	0.068	**0.061**	
FNR	0.176	0.314	0.274	0.159	0.296	**0.244**	
F-score	0.815	0.704	0.718	0.829	0.713	**0.756**	

4. Discussion

In general, TPR values for all classes are very good (Fig. 4), with an overall classification TPR of 0.76. TPR of classes 1 and 4 (rockfall/toppling and debris flow, respectively) are higher than those of classes 2, 3 and 5 (translational/rotational slide, earth flow and complex landslide, respectively). This may be due to the fact that slides, earth flows and complex landslides are more similar in shape than rockfalls and debris flows, so that the latter have more peculiar features that makes their classification easier to the ANN according to the parameters considered. This is confirmed also by the results reported in the confusion matrix in Tab. 2. Classes 2, 3 and 5 are more often confused with each other than with classes 1 and 4. This is particularly true for class 3 (otherwise 5) samples, which are wrongly predicted as classes 2 or 5 (otherwise 2 or 3) in about 2% of cases, while they are confused with class 1 or 4 in less than 1% of cases.

AUC values are higher than 0.92 for all classes, demonstrating an excellent performance of the model that is remarkably able to predict a landslide as belonging to a given class or not. In agreement with the TPR values, also the AUC values for class 1 and 4 are higher than those obtained for classes 2, 3 and 5. The model performance is particularly good if compared with that of other morphometric-based landslide classification models: Martha et al. (2010) correctly classified 69.1% of 55 landslides into five classes, while Barlow et al. (2006) achieved TPR

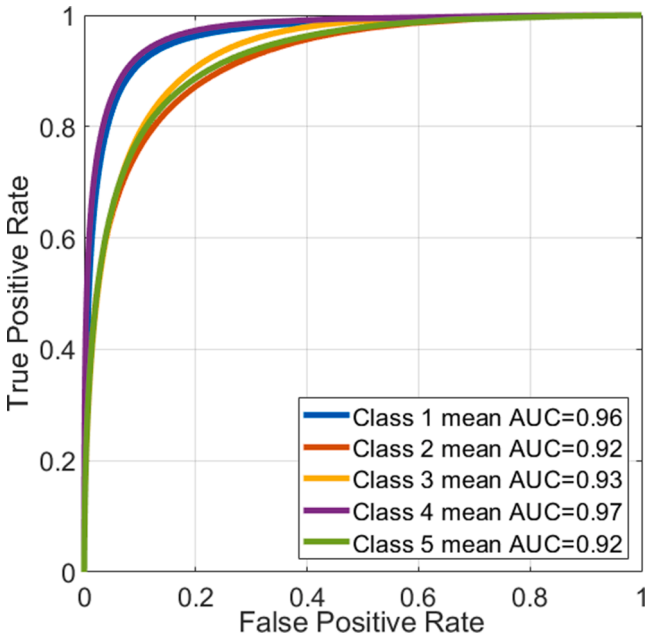


Fig. 5. ROC curves for each of the five classes.

values between 0.60 and 0.90 for a classification into three classes on a dataset with 20 samples per class. It should be also noted that the model here presented is completely data-driven and that the results were obtained without any expert-supervision.

Maps with the geographical distribution (Fig. 6) show that the model performance is very good over the entire national territory, with several, extended areas with an F-score averagely higher than 0.9. Nevertheless, the F-score distribution is not homogeneous in all the regions and presents geographical-related patterns, especially for slides, earth flows and complex landslides (class 2, class 3 and 5, respectively). The best results (F-score $>$ 0.7) for slide-type landslides are located in the central-northern regions (Tuscany, Emilia-Romagna, parts of Lombardy and Piedmont) and in southern-west part of the peninsula (Calabria). Conversely, the highest densities of low values (F-score $<$ 0.5) are located on the northeast (Friuli Venezia Giulia) and northwest (Liguria) regions as well as on the southern part of central Italy. In the two major Italian islands, Sardinia and Sicily, the number of landslides is too poor to perform reliable statistics.

F-scores of earth flows (class 3) are particularly good ($>$0.7) along a NE-verging curved band that corresponds to the Apennine mountain belt, and in Sicily region (the southern island). In the central regions, F-scores are moderately good ($>$0.5), while lower values are not numerous and randomly distributed over the national territory.

F-scores of class 5, complex landslides, are very good ($>$0.7) in the northern part of the Apennine chain and in Sicily, moderately good ($\simeq$0.6) in the northwest and in Calabria, with lower values ($<$0.5) mainly concentrated in the north-northeast and in central Italy.

Rockfalls and debris flows (classes 1 and 4) show a more homogeneous distribution and very good results: high F-scores ($>$0.7) are uniformly distributed over the entire national territory, while low values ($<$0.5) are few in number, randomly distributed and not clustered in meaningful concentrations.

The presence of clusters of low F-scores observed in some regions (Fig. 6) can be partially due to a density of landslides that is too low to calculate reliable values of spatial parameters. Contextually, the inhomogeneity of landslide mapping criteria adopted during the inventory production could have given raise to such heterogeneous distribution of the classification performance. The IFFI is produced by using different methods (aerial photo interpretation, historical documents, and field surveys) and integrates information from landslide inventories compiled

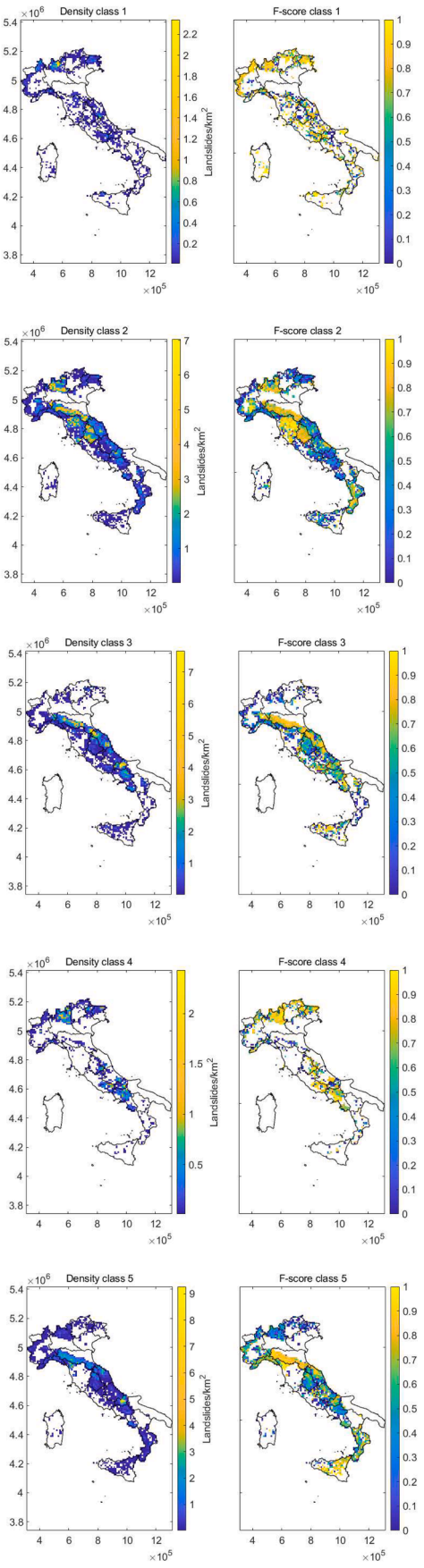


Fig. 6. Maps of landslide density and mean F-score per class per square kilometre.

by local authorities. As a consequence, the presence of areas where collected data are less accurate or underestimated has been detected in the IFFI (Trigila et al., 2010). In this context, the proposed classification method could provide an aid in validating the classification accuracy of an inventory. Actually, it should be considered that the training of the ANN is carried out by considering a national level dataset, thus optimising an overall classification model considering the full variety of classification methodologies used. In particular, it may be possible to identify areas where the performances obtained are generally lower than those expected and to investigate whether such inhomogeneities are introduced by inaccurate local classification criteria. The proposed method could be also used to identify individual landslides whose classification in the inventory might be wrong or for which the relative shape has not been accurately recorded. In this case, those landslides classified by the proposed method with a high probability in a class other than that of the catalogue could be identified for a subsequent accurate check. If the potential of the proposed method in identifying possible errors were confirmed, this could be used for a supervised or automatic correction of the database. The correction of the dataset, besides being of interest for the entities that manage it, would also allow to train the ANN with a smaller number of mislabelled samples, this potentially guaranteeing better classification performances. The presence of mislabelled samples in the training phase negatively influences the classification capacity of the ANN.

As a third application, the trained ANN can be applied to existing inventories in order to provide a landslide type classification when this information is missing.

Fig. 7 reports the results of the features importance analysis as the mean and standard deviation of the PFI score of each input parameter after 100 replicates. Parameters are sorted by a decreasing mean PFI.

PFI analysis shows that only a couple of terrain parameters is, in fact, meaningful for the classification, while vicinity and shape parameters play a key role. In particular, "Compactness IPQ" is the most important parameter. This result can be due to the fact that rockfalls and slides are generally more compact shaped than flows, which frequently show elongated shapes. This finding supports Taylor et al. (2018)'s observation on the more elliptical and compact shapes of low-mobility landslides. Among the shape-related factors, also "Compactness Moment of Inertia" and "HullPerimeter" have a major role. The latter may be seen as a measure of the indentation of the landslide shape. Intuitively, flows and complex landslides, being often sourced from different detachment points, can assume more indented shapes than rockfall and slides.

The most interesting result is the high PFI score of various spatial parameters. The number of near landslides, within a 1.5-km radius, is important to drive the classification, yet their distance from near landslides seems to be less important ("NearMeanDist" in Fig. 7), probably due to the fact that the mean distance among landslides within the given radius is relatively constant. On the other hand, the type of near landslides plays a key role in the classification: specifically, the percentage of slides, earth flows and debris flows in the vicinity deeply affects the final performance. This can be due to the fact that similar geo-environmental conditions may yield a similar distribution of these types of landslides. These can be the cases of hilly regions covered by weak lithologies and affected by numerous slides (Rosi et al., 2018) or high mountain steep slopes where several debris flows channelize into a single major valley stream (Guzzetti, 2000).

Rockfalls are also known to cluster along a common steep rock face (Varnes, 1978) and the percentage of class-1 landslides within a 1.5-km radius had considerable weight (Fig. 7) in the model results. This finding is well supported by the high F-scores assumed by the rockfall class along coastal areas and in the Alpine region, as shown in Fig. 6. The percentage of complex landslides within 1.5-km radius resulted to have the lowest importance, among the spatial parameters, for the classification success.

Finally, the TRI and the standard deviation of elevation had high PFI scores among the terrain parameters. This can be explained by the fact that debris flows can cause the outcropping of irregular bare rock surfaces along their run-out channel, due to their erosive capability, while earth flows can give rise to heterogeneous terrain deformation, due to different activity rates within the landslide body. Conversely, slides, usually moving along pre-existing or regular surfaces (e.g. strata planes), may cause lower terrain roughness.

The mean value of slope gradient of the landslide also resulted to be a relevant terrain factor: rockfalls and debris flows are often characterised by high slope angles in their detachment and run-out zones while slides and earth flows may develop even in hilly and more gentle slopes composed by soft rocks.

At last, we grouped all the parameters in four groups: (1) North and East coordinates of the landslide polygon centroids, (2) shape-related, (3) terrain and (4) spatial parameters. First, we run the model adopting all the possible combinations of these groups and then we plotted the results for each possible combination. Fig. 8 shows the AUC of every class changes depending on the different combinations.

We can observe that the classification of rockfall and debris flow

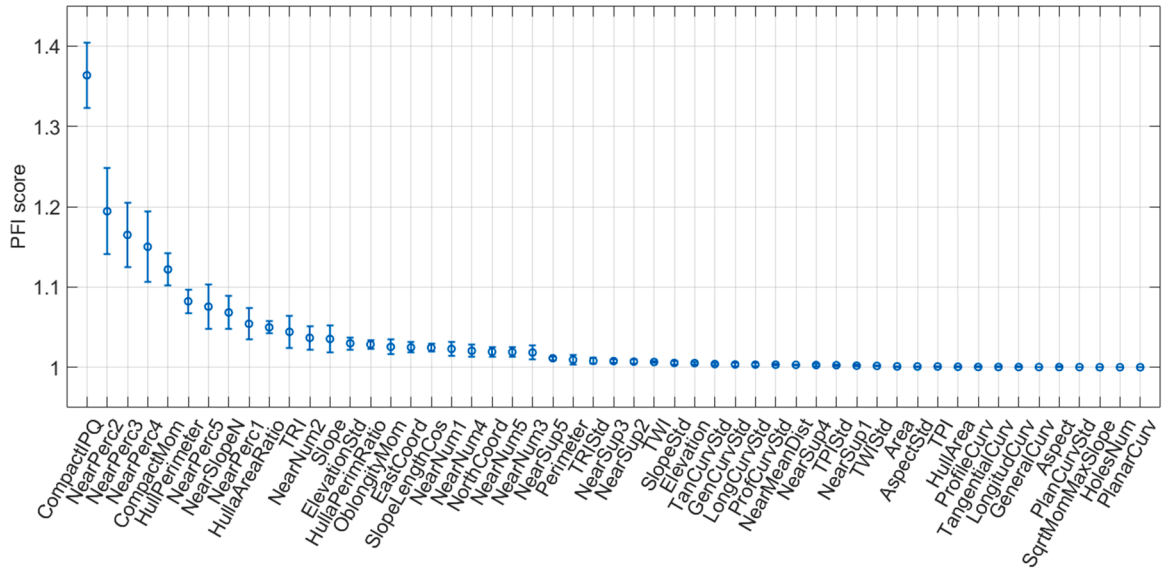


Fig. 7. Output of the Permutation Feature Importance analysis.

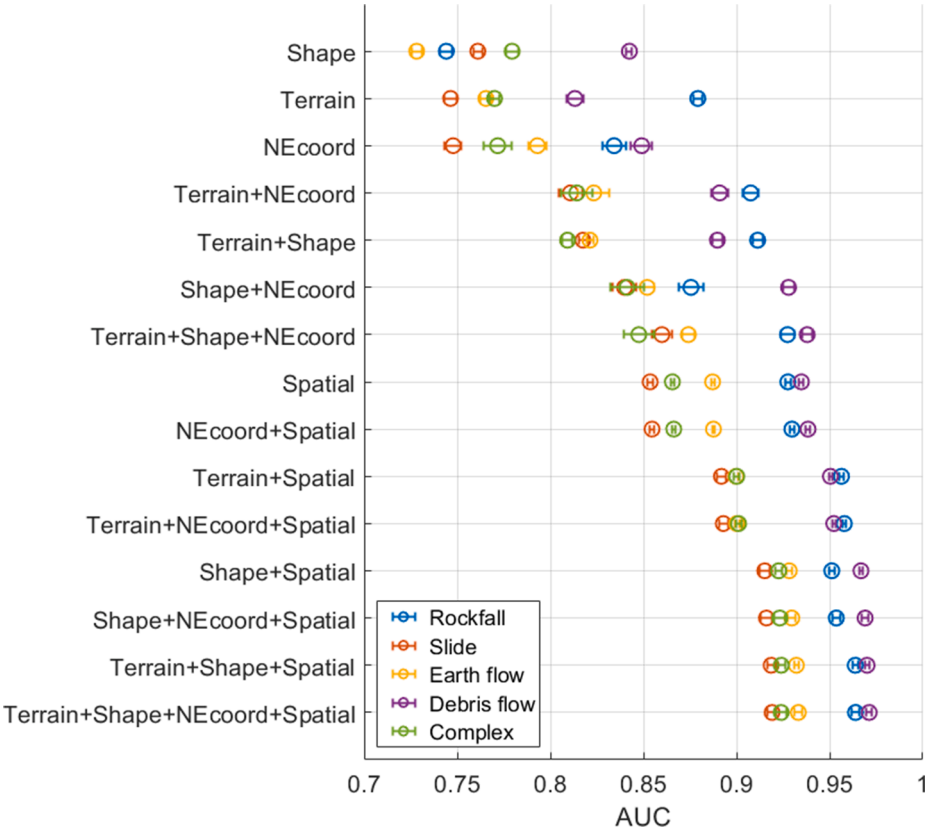


Fig. 8. Effect of parameters combinations on the classification performance.

generally reaches the best results, regardless of the parameters combination. Nevertheless, the choice of the parameters deeply affects the final performances. In this regard, it is clear that using only shape-related or terrain parameters is not sufficient to achieve a reliable classification of landslide types. In particular, the classification metrics of rockfall/toppling and translational/rotational slide improves considerably if terrain parameters are considered. The same improvement is observed for debris flow if shape factors are considered. This analysis confirms the key role played by the spatial parameters in the improvement of the classification performance. From Fig. 8 we can infer that, rather than the North and East coordinates of a landslide, the number and the type of the failures located in its vicinity play a crucial role. These parameters are rarely considered in landslide geostatistical models: with these findings, we aim to provide an original insight for discussion within the community. In this regard, the assumption that the conditions that caused a landslide in the past can cause a similar event also in the future is largely accepted by the community (Guzzetti et al., 1999). Therefore, according to these findings, we can assume that these conditions may occur again not only in time, but also in space and close to the location of occurrence of a landslide in the past. Under this assumption, we could infer that a slope failure may tend to be surrounded by failures of the same type and, thus, this information may represent a meaningful parameter to be modelled in landslide geostatistics.

5. Conclusions

In this work, we presented an innovative data-driven model of landslide type classification and its application to a national-scale dataset of landslides in Italy. The method relies on the use of ANN and of input parameters that can be easily extracted from polygonal landslides inventories and DEM. The results achieved a very good performance in landslide classification (TPR of 0.76 for a five-class overall classification of over 275,000 landslides). They provide insight on the meaningfulness of the landslide spatial distribution for geostatistical modelling and foster further investigation on the link between landslide process and its shape.

The method can be applied to any polygonal inventory, like those produced by automatic mapping procedures from Earth Observation imagery, in order to automatically define the types of landslides when this information is missing. It should be noted that the experiment presented here can be reproduced in other geographical areas where there is the availability of a DEM with spatial resolution similar to the one used in this work and of a polygonal landslide catalogue that includes the landslide types modelled here. Actually, the ANN does not require the input of site-specific information (such as geology or land cover) in order to perform the landslide type classification. Finally, this paper can represent a step forward to the standardisation of landslide inventories compiled from different sources, producing tangible effects for hazard modelling applications.

Funding

This work was partially supported by the Italian Ministry of Research and Education [OT4CLIMA project under the PON–FSC programme];

CRediT authorship contribution statement

Gabriele Amato: Conceptualization, Investigation, Methodology, Resources, Validation, Data curation, Visualization, Writing – original draft. **Lorenzo Palombi:** Conceptualization, Methodology, Resources, Data curation, Formal analysis, Software, Validation, Visualization, Writing – original draft. **Valentina Raimondi:** Validation, Funding acquisition, Project administration, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Amato, G., Eisank, C., Castro-Camilo, D., Lombardo, L., 2019. Accounting for covariate distributions in slope-unit-based landslide susceptibility models. A case study in the alpine environment. Eng. Geol. 260, 105237.

Ayalew, L., Yamagishi, H., Ugawa, N., 2004. Landslide susceptibility mapping using GIS-based weighted linear combination, the case in Tsugawa area of Agano River, Niigata Prefecture, Japan. Landslides 1 (1), 73–81.

Bachi, R., 1999. New methods of geostatistical analysis and graphical presentation: distributions of populations over territories. Springer Science and Business Media. Published in cooperation with The Israel Academy of Sciences and Humanities. ISBN 0-306-45544-7.

Barlow, J., Franklin, S., Martin, Y., 2006. High spatial resolution satellite imagery, DEM derivatives, and image segmentation for the detection of mass wasting processes. Photogramm. Eng. Remote Sens. 72 (6), 687–692.

Beven, K.J., Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. Hydrol. Sci. J. 24 (1), 43–69.

Crozier, M.J., 2005. Multiple-occurrence regional landslide events in New Zealand: hazard management issues. Landslides 2 (4), 247–256.

Cruden, D.M., Varnes, D.J., 1996. Landslides: investigation and mitigation. Chapter 3-Landslide types and processes. Transportation Research Board Special Report, p. 247.

Dai, F.C., Lee, C.F., 2002. Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. Geomorphology 42 (3-4), 213–228.

De Veaux, R.D., Ungar, L.H., 1994. Multicollinearity: A tale of two nonparametric regressions. In: Selecting Models from Data. Springer, New York, NY, pp. 393–402.

Dikau, R., Cavallin, A., Jäger, S., 1996. Databases and GIS for landslide research in Europe. Geomorphology 15 (3–4), 227–239.

Evans, I.S., 1980. An integrated system of terrain analysis and slope mapping. Zeitschrift für Geomorphologie. Supplementband Stuttgart 36, 274–295.

Garg, A., Tai, K., 2012. Comparison of regression analysis, artificial neural network and genetic programming in handling the multicollinearity problem. In: 2012 Proceedings of International Conference on Modelling, Identification and Control, pp. 353–358.

Gariano, S.L., Guzzetti, F., 2016. Landslides in a changing climate. Earth Sci. Rev. 162, 227–252.

Guzzetti, F., 2000. Landslide fatalities and the evaluation of landslide risk in Italy. Eng. Geol. 58 (2), 89–107.

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 (1-4), 181–216.

Guzzetti, F., Mondini, A.C., Cardinali, M., Fiorucci, F., Santangelo, M., Chang, K.-T., 2012. Landslide inventory maps: New tools for an old problem. Earth Sci. Rev. 112 (1-2), 42–66.

Heerdegen, R.G., Beran, M.A., 1982. Quantifying source areas through land surface curvature and shape. J. Hydrol. 57 (3-4), 359–373.

Hungr, O., Evans, S.G., Bovis, M.J., Hutchinson, J.N., 2001. A review of the classification of landslides of the flow type. Environmental and Engineering Geoscience 7 (3), 221–238.

Hungr, O., Leroueil, S., Picarelli, L., 2014. The Varnes classification of landslide types, an update. Landslides 11 (2), 167–194.

Hutchinson, J.N., 1988. Morphological and geotechnical parameters of landslides in relation to geology and hydrogeology. General report. Landslides. Proc. 5th ISL, Lausanne. Balkema, Rotterdam, 1. C. Bonnard, pp. 3–35.

Johnson, J.M., Khoshgoftaar, T.M., 2019. Survey on deep learning with class imbalance. J. Big Data 6, 27. https://doi.org/10.1186/s40537-019-0192-5.

Li, Y., Mo, P., 2019. A unified landslide classification system for loess slopes: A critical review. Geomorphology 340, 67–83.

Mantovani, F., Soeters, R., Van Westen, C.J., 1996. Remote sensing techniques for landslide studies and hazard zonation in Europe. Geomorphology 15 (3-4), 213–225.

Martha, T.R., Kerle, N., Jetten, V., van Westen, C.J., Kumar, K.V., 2010. Characterising spectral, spatial and morphometric properties of landslides for semi-automatic detection using object-oriented methods. Geomorphology 116 (1-2), 24–36.

Martha, T.R., Kerle, N., van Westen, C.J., Jetten, V., Kumar, K.V., 2011. Segment optimization and data-driven thresholding for knowledge-based landslide detection by object-based image analysis. IEEE Trans. Geosci. Remote Sens. 49 (12), 4928–4943.

Mezaal, M.R., Pradhan, B., 2018. An improved algorithm for identifying shallow and deep-seated landslides in dense tropical forest from airborne laser scanning data. Catena 167, 147–159.

Moosavi, V., Talebi, A., Shirmohammadi, B., 2014. Producing a landslide inventory map using pixel-based and object-oriented approaches optimized by Taguchi method. Geomorphology 204, 646–656.

Osserman, R., 1978. The isoperimetric inequality. Bull. Am. Math. Society 84 (6), 1182–1239.

Popescu, M., 2001. A suggested method for reporting landslide remedial measures. Bull. Eng. Geol. Environ. 60 (1), 69–74.

Pourghasemi, H.R., Gayen, A., Park, S., Lee, C.W., Lee, S., 2018. Assessment of landslide-prone areas and their zonation using logistic regression, logitboost, and naïvebayes machine-learning algorithms. Sustainability 10 (10), 3697.

Prakash, N., Manconi, A., Loew, S., 2020. Mapping Landslides on EO Data: Performance of Deep Learning Models vs Traditional Machine Learning Models. Remote Sens. 12 (3), 346.

Putin, E., Mamoshina, P., Aliper, A., Korzinkin, M., Moskalev, A., Kolosov, A., Ostrovskiy, A., Cantor, C., Vijg, J., Zhavoronkov, A., 2016. Deep biomarkers of human aging: Application of deep neural networks to biomarker development. Aging (Albany NY) 8, 1021–1033.

Riley, S.J., DeGloria, S.D., Elliot, R., 1999. Index that quantifies topographic heterogeneity intermountain. J. Sci. 5 (1–4), 23–27.

Rosi, A., Tofani, V., Tanteri, L., Tacconi Stefanelli, C., Agostini, A., Catani, F., Casagli, N., 2018. The new landslide inventory of Tuscany (Italy) updated with PS-InSAR: geomorphological features and landslide distribution. Landslides 15 (1), 5–19.

Taalab, K., Cheng, T., Zhang, Y., 2018. Mapping landslide susceptibility and types using random forest. Big Earth Data 2, 159–178.

Taylor, F.E., Malamud, B.D., Witt, A., Guzzetti, F., 2018. Landslide shape, ellipticity and length-to-width ratios. Earth Surf. Proc. Land. 43 (15), 3164–3189.

Tharwat, A., 2021. Classification assessment methods. Applied Compting and Informatics 17 (1), 168–192.

Trigila, A., Iadanza, C., Spizzichino, D., 2010. Quality assessment of the Italian Landslide Inventory using GIS processing. Landslides 7 (4), 455–470.

Van Hulse J, Khoshgoftaar TM, & Napolitano A. (2007). Experimental perspectives on learning from imbalanced data. In: Proceedings of the 24th international conference on machine learning. ICML '07. ACM, New York, NY, USA. 2007. p. 935–942. 10.1145/1273496.1273614.

Varnes, D.J., 1978. Slope movement types and processes. Special Rep. 176, 11–33.

Wood, J.L., Harrison, S., Reinhardt, L., Taylor, F.E., 2020. Landslide databases for climate change detection and attribution. Geomorphology 355, 107061. https://doi.org/10.1016/j.geomorph.2020.107061.

Zevenbergen, L.W., Thorne, C.R., 1987. Quantitative analysis of land surface topography. Earth Surf. Proc. Land. 12 (1), 47–56.