Application of neural network to assess landslide hazard and comparison with bivariate and multivariate statistical analyses

Rudolf Tornyai[1], Martin Bednarik[1] & Aleš Havlín[2]

[1]Department of Engineering Geology, Faculty of Natural Sciences, Comenius University in Bratislava, Bratislava, Mlynská dolina, Ilkovičova 6; tornyai@fns.uniba.sk
[2]Czech Geological Survey, Branch Brno, Leitnerova 22, 658 69 Brno, Czech Republic

AGEOS

Aplikácia neurónovej siete pri hodnotení zosuvného hazardu a porovnanie s bivariačnou a multivariačnou štatistickou analýzou

Abstract: Landslide hazard in the Žilina area in northern Slovakia is assessed using neural network analysis. Four input parameters are evaluated, they are presented as a result of statistical processing in the form of parametric maps. Statistical evaluation was executed in ArcGIS environment; neural network was calculated in Matlab. The output of this study is a prognostic landslide hazard map. Further, the result was compared with the hazard map created using bivariate and multivariate statistical analyses through ROC curves. Area below curve (AUC) calculated from ROC curve shows accuracy of individual models. It can be stated that the NN´s AUC is equal to 0.924, what represents the rate of success 92.4%; bivariate multivariate analyses AUC is equal to 0.852 and 0.919.

Key words: landslide hazard assessment, neural network, bivariate analysis, multivariate conditional analysis, Žilina region, ROC curves

1. INTRODUCTION

Nowadays, due to climate change, heavy rainfalls are becoming more frequent, which causes the landslides occur more often. In many parts of the world they represent significant environmental threat. In Slovakia, there are many works published recently regarding landslide activity, e.g. Liščák et al. (2010), Dostál et al. (2014), Šilhán et al. (2014), or Putiška et al. (2015). That is why we experience the full development of methods that can predict where the landslide could occur in the future. The most important step is to analyze the factors of their origin. The aim of efforts of many experts is to create the forecasts, according to which we would be able to predict the origin and development of landslides in the future with a high probability. One of the last proven methods is assessment of the landslide hazard using Neural Networks. The theoretical base of this method is sufficiently discussed in the literature. The Neural Networks are a computational method of data analysis. It is an extension of traditional statistical methods such as regression (White, 1989), and the approximation function (Baum & Haussler, 1989; Hertz et al., 1991).

Neural Networks (NN) is a massive parallel tool that tends to preserve experimental knowledge and its further use. It mimics the human brain in two aspects: 1. knowledge is collected during learning; 2. connections between neurons (synaptic weights - SV) are used to store the knowledge. This is one of the NN definitions accepted by NN community and was inspired from biological systems. Crudely saying it is brain simulation. On first impression, this highly abstract discipline finds many applications in practice and becomes a tool for solving problems in a wide range of professional uses. One of the most important features of the NN is that it works as a universal functions approximator. This approach is useful for systems with extremely complex, or almost impossible description. In such a situation when we have data that enter the system, and the outputs corresponding to them, then we can use a suitable NN and try to teach it to behave like system using training data (mentioned inputs and outputs). This is a very important point, as it determines the application of NN in practice (Hertz et al., 1991). It is possible to find out a sufficient number of examples where NN was used to evaluate landslide hazards in literature. For instance, Aleotti et al. (1998) applied NN in northern Italy for the classification of landslides according to the degree of hazard. Macchi & Deravignone (2006) published study which described the implementation of the methodology of neural networks in the GIS environment. For the purpose of artificial neural networks application it was necessary to develop specific software applications. The aim was to create a kind of "bridge" between GIS platforms and simulators of NN. Ermini et al. (2005) used an artificial neural network for landslide hazard assessment. They used two types of neural networks: a Multi-Layer Perceptron (MLP) and probabilistic neural network (PNN). Pradhan & Lee (2009) in Malaysia (Penang) applied NN for assessing landslide risk, when feed-forward neural network was used with back propagation error learning algorithm. Results of study were maps of landslide hazard degree determination. Nowadays, the possibility of using computing technologies experiences full development, together with increasing use of geographic information systems (GIS), which are inevitable in the application of NN for assessing landslide hazard.

The Žilina region (Fig. 1) is located in the northern part of Slovakia. Geologically, the area is composed of (from the northwest to the southeast) Silesian, Magura, and Oravic units. The study was focused predominantly to the Magura Unit which

Manuscript received 2016-03-16
Revised version accepted 2016-05-02

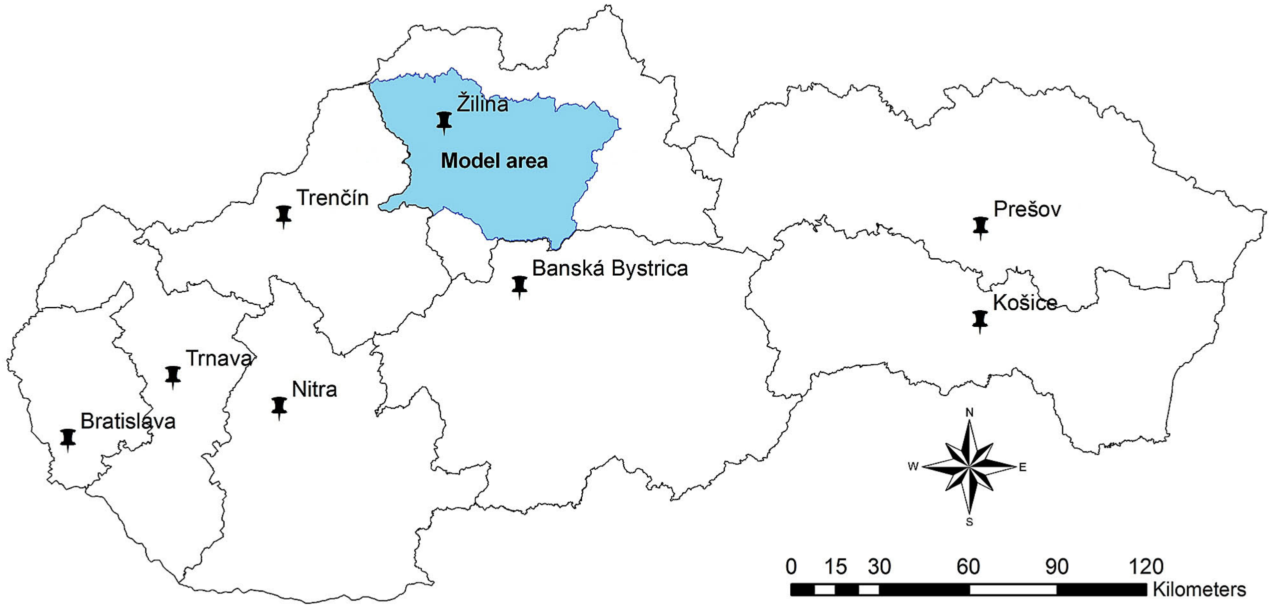

Fig. 1: Geographic location of the study area

contains the Rača, Bystrica, and Krynica subunits and forms about 1 km wide strip along the northern edge of the Pieniny Klippen Belt (Potfaj et al. 2003).

2. APPLICATION OF NEURAL NETWORK

Application of NN for remote sensing (RS) data interpretation was motivated by the ability to efficiently handle very large amounts of data from various sources. NN in the simplest sense transform input to output, thus it belongs to the same class of techniques such as automatic recognition of symptoms, regression, spectral and textural classification. Due to importance of these techniques the increasing tendency of using NN in RS is not a surprise.

The rapid increase of NN applications in RS is mainly a consequence of their skills: working more accurately and much faster than other techniques like e.g. statistical classifiers, especially when the feature space and complex data sources have different statistical distribution. Thus it is clear that one of the main advantages of the NN is the ability to effectively handle the large amount of RS data.

For landslide hazard assessment multilayer forward neural network, the so-called Multi-Layer Perceptron (MLP) and the learning algorithm of back propagation error were used, approach by Pradhan & Lee (2009). The MLP, as the name implies, consists of a series of layers, each consisting of a set of nodes (neurons). Within the feed-forward neural network only forward connections between neurons exist. Each neuron of one layer sends signals to each neuron of the next layer. Connections to the previous layer do not exist.

A neural network consists of a number of interconnected nodes. Each node is a simple processing element that responds to the weighted inputs it receives from other nodes. The arrangement of the nodes is called network architecture (Fig. 2). The receiving node sums the weighted signals from all the nodes that were connected with in the preceding layer. Formally, the input that a single node receives, is weighted according to equation (1):

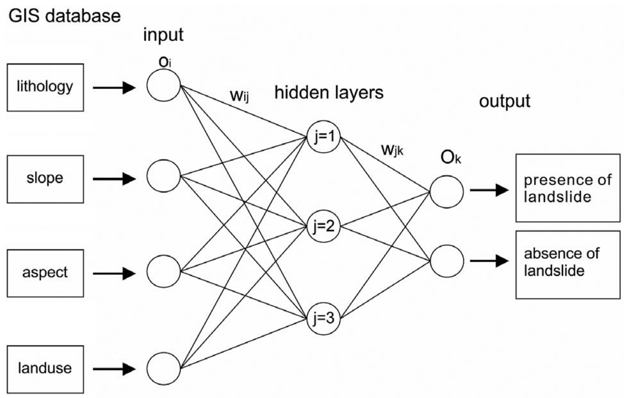

Fig. 2: Basic architecture of neural network

$$net_j = \sum_i w_{ij} \times o_i \quad (1)$$

where w_{ij} represents the weights between nodes i and j, and o_i is the output from node j, according to:

$$o_j = f(net_j) \quad (2)$$

The function f is usually a non-linear sigmoid function that is applied to the weighted sum of inputs before the signal propagates to the next layer. One advantage of a sigmoid function is that its derivative can be expressed in terms of the function itself:

$$f'(net_j) = f(net_j)(1 - f(net_j)) \quad (3)$$

The network used in this study consisted of three layers. The first layer is the input layer, where the nodes were the elements of a feature vector. The second layer is the internal or "hidden" layer. The third layer is the output layer that presents the output data. Each node in the hidden layer is interconnected with the nodes in both the preceding and following layers by weighted connections (Atkinson & Tatnall, 1997). The error, E, for an input training pattern, t, is a function of the desired output.

$$E = \frac{1}{2}\sum_k (d_k - o_k) \tag{4}$$

The error is propagated back through the neural network and is minimized by adjusting the weights between layers. The weight adjustment is expressed as:

$$w_{ij}(n+1) = \eta(\delta_j \times o_i) + \alpha \Delta w_{ij} \tag{5}$$

Where η is the learning rate parameter (set to $\eta = 0.01$ in this study), $\delta\, j$ is an index of the rate of change of the error, and α is the momentum parameter (set to $\alpha = 0.01$ in this study). The factor δ_j is dependent on the layer type. For example, for hidden layers:

$$\delta_j = (\sum \delta_k w_{jk}) f'(net_j) \tag{6}$$

and for output layers:

$$\delta_j = (d_k - o_k) f'(net_k) \tag{7}$$

This process of the feed-forward signals and the back-propagating error is repeated iteratively until the error of the network as a whole is minimized or reaches an acceptable magnitude. Using the back-propagation training algorithm, the weights of each factor can be determined and may be used for classification of data (input vectors) that the network had not seen before. Zhou (1999) described a method for determining the weights using back-propagation. From equation (2), the effect of an output, o_j, from a hidden layer node, j, on the output, o_k, from an output layer (node k) can be represented by the partial derivative of o_k with respect to o_j as:

$$\frac{\partial o_k}{\partial o_j} = f'(net_k) \times \frac{\partial (net_k)}{\partial o_j} = f'(net_k) \times w_{jk} \tag{8}$$

Equation (8) produces both positive and negative values. If the effect's magnitude is all that is of interest, then the importance (weight) of node j relative to another node j_0 in the hidden layer may be calculated as the ratio of the absolute values derived from equation (8):

$$\frac{|\partial o_k|}{|\partial o_j|} / \frac{|\partial o_k|}{|\partial o_{j0}|} = \frac{|f'(net_k) \times w_{jk}|}{|f'(net_k) \times w_{j0k}|} = \frac{|w_{jk}|}{|w_{j0k}|} \tag{9}$$

We should mention that w_{j0k} is simply another weight in w_{jk} other than w_{ik}. For a given node in the output layer, the results of equation (9) show that the relative importance of a node in the hidden layer is proportional to the absolute value of the weight connecting the node to the output layer. When the network consists of output layers with more than one node, then equation (9) cannot be used to compare the importance of two nodes in the hidden layer:

$$w_{j0k} = \frac{1}{J} \times \sum_{j=1}^{J} |w_{jk}| \tag{10}$$

$$t_{jk} = \frac{|w_{jk}|}{\frac{1}{J} \times \sum_{j=1}^{J} |w_{jk}|} = \frac{J \times |w_{jk}|}{\sum_{j=1}^{J} |w_{jk}|} \tag{11}$$

Therefore, with respect to node k, each node in the hidden layer has a value that is greater or smaller than unity, depending on whether it is more or less important, respectively, than an average value. All the nodes in the hidden layer have a total importance with respect to the same node, given by:

$$\sum_{j=1}^{J} t_{jk} = J \tag{12}$$

Consequently, the overall importance of node j with respect to all the nodes in the output layer can be calculated by:

$$t_j = \frac{1}{K} \times \sum_{j=1}^{K} t_{jk} \tag{13}$$

Similarly, with respect to node j in the hidden layer, the normalized importance of node j in the input layer can be defined by:

$$s_{ij} = \frac{|\omega_{ij}|}{\frac{1}{I} \times \sum_{i=1}^{I} |\omega_{ij}|} = \frac{I \times |\omega_{ij}|}{\sum_{i=1}^{I} |\omega_{ij}|} \tag{14}$$

The overall importance of node i with respect to the hidden layer is:

$$s_i = \frac{1}{J} \times \sum_{j=1}^{J} s_{ij} \tag{15}$$

Correspondingly, the overall importance of input node i with respect to output node k is given by (Pradhan & Lee, 2009):

$$st_i = \frac{1}{J} \times \sum_{j=1}^{J} s_{ij} \times t_j \tag{16}$$

Neural network gets its ability to transform input data into output during the learning process. Learning can be defined as a process in which the parameters of neural networks (w_{ij} and w_{jk} - synaptic weights) vary based on some rules. The nature of these rules, the effect of altering synaptic weights, determines the type of learning. Under the learning of NN we understand adaptation of NN that in the end of the learning will be the bearer of knowledge gained during learning. Learning is a fundamental and essential characteristic of neural networks. As already mentioned, the most used learning algorithm for multilayered NN is a method of back-propagation error (back-propagation training algorithm), whose authors are Rumelhart et al. (1986). The algorithm consists of three stages: the feed-forward spread signal between neurons connections, back-propagation of error and modification of synaptic weights so as to minimize the amount of error between the desired and actual output of the network. Neural network simulation took place in the neural networks module in the environment of the Matlab software package.

To verify the results of Artificial Neural Network (**ANN**) application and comparison with two other statistical methods are used: bivariate and multivariate which are constructed as in the study of Tornyai & Dunčko (2013).

The methodology of landslide hazard assessment using statistical methods in a GIS environment is based on an appropriate choice of the factors affecting the stability of slopes. Statistical processing of landslide hazard assessment is based on the geological principle of actualism that landslides will occur in places where they occurred in the past respectively in present under the similar activation conditions.

Selected factors, which influenced slope movements, are processed into a form of parametric maps and like this they are entering into the process of statistical evaluation using map algebra in GIS/Matlab environment.

According to the chosen statistical method a comparison of parametric maps factors with the landslide inventory map of model area follows. Conclusions resulting from statistical comparisons are extrapolated to the whole area of the region and the result is a prognostic hazard map.

3. INPUT PARAMETERS

Model area of Žilina region has a total area of 3132.2 square kilometers. The region was just administratively allocated. Four input factors that most influence the instability of slopes were evaluated. Each of the factors was processed into a form of parametric map in ArcGIS environment. Subsequently, the input parameters were extracted into the TIFF format and entered MATLAB environment. Evaluated parameters are: geological conditions, slope angle, slope aspect, the current land-use, and the registered slope deformations. Parametric maps were constructed from four vector maps processed at a scale of 1:10 000.

3.1 Geological conditions

Geological setting of the area is one of the most important factors affecting the formation and evolution of slope deformations. Geological structure is characterized by rocks of the Magura Group of the Outer Western Carpathians and includes Rača Unit (north), Biele Karpaty Unit and Bystrica - Orava Unit (in the south). In

Tab. 1 Spatial distribution of lithological units

Category	Genetic type	Lithological characteristics	Area [km^2]	Area [%]
1	anthropogeneous sediments	fills, heaps and dumps	1.87	0.06
2	fluvial sediments	alluvial soils, sandy to gravelly soils of valley and mountain streams	393.58	12.57
3	proluvial sediments	clay, loam, gravels of proluvial fans	50.87	1.62
4	slope sediments	lithofacial undifferentiated deluvium and debris	455.93	14.56
5	chemogene-organogenic sediments	freshwater limestone, travertine, tufa, calcareous sinter	2.30	0.07
6	organic sediments	peat, peat humus clay	1.79	0.06
7	Neogene deposits	compacted gravel	29.10	0.93
8	carbonates of Central Carpathian Palaeogene Basin	breccias, conglomerates, sandstones, siltstones, limestones	0.12	0.00
9	siliciclastic deposits of Central Carpathian Palaeogene Basin	claystones prevail over sandstones and conglomerates	91.08	2.91
10	siliciclastic deposits of Outer Western Carpathians	claystones, siltstones, sandstones	559.66	17.87
11	Pieniny Klippen Belt	dark claystone, marl, spotted marl and limestone marl	253.87	8.11
12	limestones and dolomites of the Tatric, Fatric, and Hronic units	limestones, dolomites	1004.88	32.08
13	siliciclastic deposits of Tatric Unit	red sandstone, sandy dolomite and clay-slate, conglomerates, feldspathic offal, arkose, suborder sandy shale	4.37	0.14
14	Tatric crystalline basement	granites, granodiorites, paragneiss	282.86	9.03

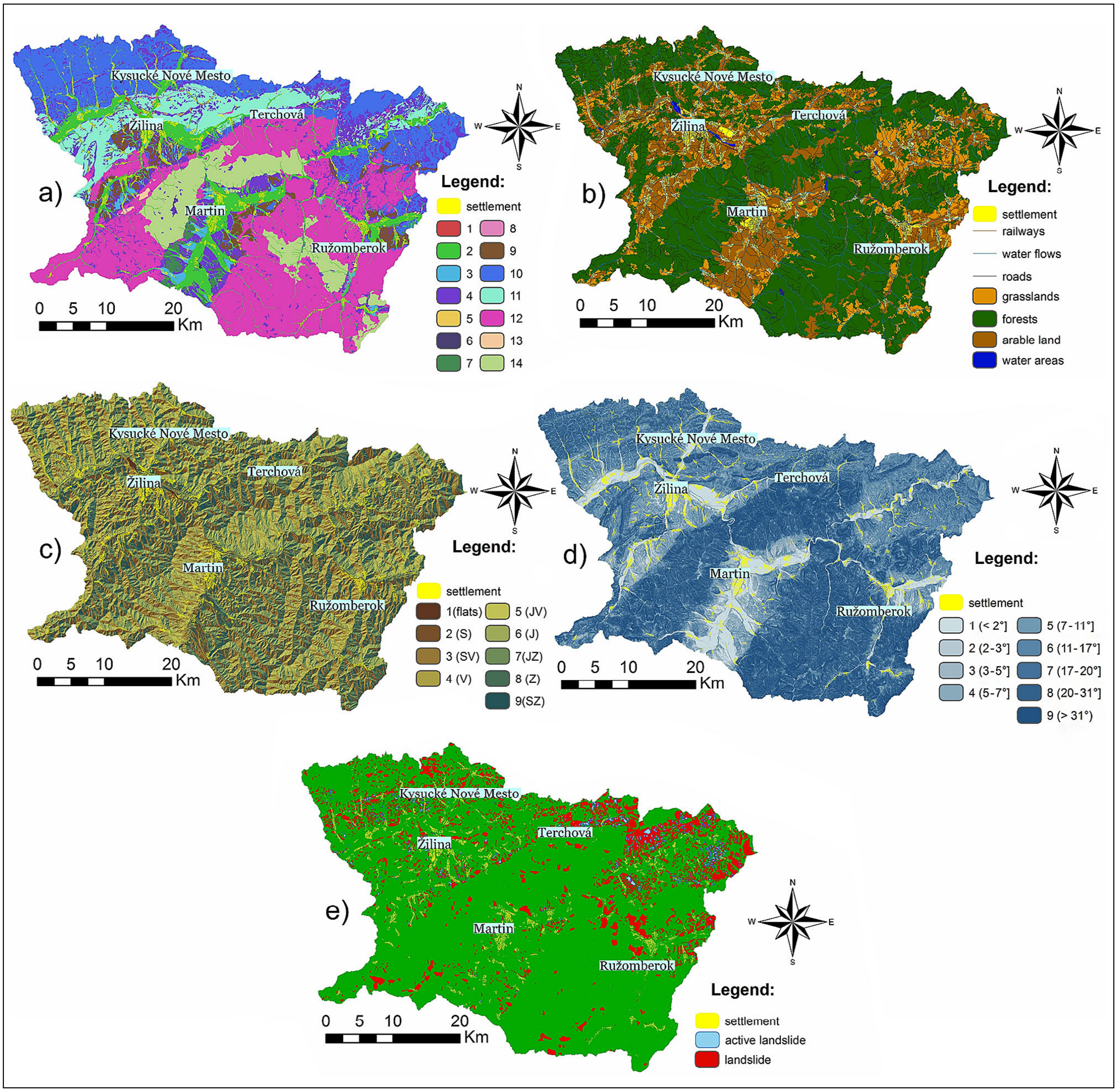

Fig. 3: Input parametric maps

a) map of lithological units (1 – anthropogeneous sediments; 2 – fluvial sediments; 3 – proluvial sediments; 4 – slope sediments; 5 – chemogenic-organogenic sediments; 6 - organic sediments; 7 – Neogene deposits; 8 – carbonates of Central Carpathian Palaeogene Basin; 9 – siliciclastic deposits of Central Carpathian Palaeogene Basin; 10 – siliciclastic deposits of Outer Western Carpathians; 11 – Pieniny Klippen Belt; 12 -limestones and dolomites of the Tatric, Fatric, and Hronic units; 13 – siliciclastic deposits of Tatric Unit; 14 – Tatric crystalline basement); b) landuse map; c) aspect; d) slope angle map; e) landslide inventory map

the Rača Unit (from bottom to top) Beloveža Formation (mainly claystones) overlies Soláň Formation (mainly sandstones) and is overlain by Zlín Formation (mostly claystones). The Bystrica Unit has similar stratigraphic sequence, except of the uppermost member - Bystrica Member containing more claystones than the Zlín Member in the Rača Unit.

Physical and mechanical characteristics due to the lithological composition of the rock environment are an important factor affecting the stability of the slope. Reclassified map of engineering geological conditions contains 14 lithological units (1 – anthropogeneous sediments; 2 – fluvial sediments; 3 – proluvial sediments; 4 – slope sediments; 5 – chemogenic-organogenic sediments; 6 - organic sediments; 7 – Neogene deposits; 8 – carbonates of Central Carpathian Palaeogene Basin; 9 – siliciclastic deposits of Central Carpathian Palaeogene Basin; 10 – siliciclastic deposits of Outer Western Carpathians; 11 –Pieniny Klippen Belt; 12 -limestones and dolomites of the Tatric, Fatric, and Hronic units; 13 – siliciclastic deposits of Tatric Unit; 14 – Tatric crystalline basement; Fig. 3a). The cell size is 10×10 m.

The most widespread dissemination of the model area (32%) represent limestones and dolomites. Slope sediments, which can be described as the most susceptible to formation of slope deformations represent more than 14% of the model area (Tab. 1).

3.2 Current landscape structure

The current landscape structure reflects current land use, including vegetation cover. This parameter is very dynamic, subject to relatively rapid changes in time, and therefore it is necessary to use the most up to date information during processing. Clearly most reliable sources are up to date aerial and satellite images, respectively orthophoto maps of model area. For this purpose Google maps were used and manually converted to landscape structure.

In the model area 8 elements are selected (Fig. 3b) of the current landscape structure: (1) road network, (2) settlement, (3) meadows, (4) arable land, (5) forests, (6) gardens, (7) rivers and (8) railway network. Spatial distribution of elements of the current landscape structure shows that the largest part of the model area is covered by forests (60.78%), meadows (19.21%) and arable land (12.36%) (Tab. 2).

3.3 Aspect map

Slope aspect factor is often taken into account in relation with weather and meteorological conditions in the study area. These represent e. g. prevailing wind direction, which together with the cumulated sunlight significantly affects evapotranspiration, soil moisture status and so on (Pauditš, 2005).

Parametric map of slope aspect (Fig. 3c) presents a continuous data field indicating the value of a certain angle from one of the cardinal (mostly from the north), in a counter-clockwise. Slope aspect was reclassified to 8 semiquadrants. Class nine represents a territory without reference to the cardinal points, planar (flats) areas.

In the model area slope aspect is approximately equally distributed (Tab. 3). We can see very little superiority of slopes facing northwest.

3.4 Slope angle map

Slope angle is one of the most important morphometric parameters affecting slope instability. Slope angles are expressed in degrees (range 0 – 90°, Fig. 3d). The slope angle in combination with other is a parameter significantly influencing the slope stability conditions.

For classification of slopes a methodology by Hrašna (1986) was used, which is usually used in the engineering geological mapping. Slopes are divided into nine categories shown in Tab. 4 in column interval.

More than 28% of the area is formed by slopes from 20 to 31 degrees. The second most common category are slopes with 11 to 17 degrees of inclination and makes up more than 20 % of the territory (Tab. 4). Slopes, which can be considered as highly susceptible to landslides (categories 4, 5, 6 and 7) cover more than 46 % of the area.

3.5 Landslide inventory map

In the model area of the Žilina region map of slope deformations represents binary dependent (dichotomic) variable which in the process of statistical analysis all input parametric maps are compared with. Binary raster map of slope deformations contains only Boolean values 0 and 1 (false / true), where the value 1 represents the existence of a landslide in the cell grid and the absence of landslide represents the value 0. Landslides were vectorized as polygons without distinction of main scarps which were not identified in originally used source, so as whole landslide bodies (Bednarik & Pauditš, 2010).

Currently 11.96% of area is affected by slope deformations, representing an area of 374.92 square kilometers (Fig. 3e). For the calculation of landslide hazard active landslides were used, which cover an area of 22.37 square kilometers (Grman et al., 2011).

4. RESULTS AND DISCUSSION

After processing all input parametric maps, they were exported to tiff raster format. In this format, they entered MATLAB environment. The training and testing group was selected randomly. The training sample was made up of inputs and to them pertinent outputs. One subset of the data set contains the combination where the result is always a value of 1, present landslide (active landslides), while the second part contains a

Tab. 2 Spatial distribution of present land-use

Category	Area [km^2]	Area [%]
1 (road network)	37.59	1.20
2 (settlement)	186.06	5.94
3 (meadows)	601.71	19.21
4 (arable land)	387.15	12.36
5 (forests)	1903.79	60.78
6 (gardens)	0.63	0.02
7 (rivers)	10.65	0.34
8 (railway network)	4.70	0.15

Tab. 3 Spatial distribution of slope orientation

Category	Interval	Area [km^2]	Area [%]
1 (flat)	(–1)	4.65	0.15
2 (N)	(0–22.5), (337.5–360)	401.98	12.83
3 (NE)	(22.5–67.5)	402.48	12.85
4 (E)	(67.5–112.5)	377.09	12.04
5 (SE)	(112.5–157.5)	362.80	11.58
6 (S)	(157.5–202.5)	383.59	12.25
7(SW)	(202.5–247.5)	400.22	12.78
8 (W)	(247.5–292.5)	391.39	12.50
9 (NW)	(292.5–337.5)	408.06	13.03

set of data with the result of 0, no landslide area (stable area). Using the back propagation algorithm neural network in MATLAB was trained.

Artificial neural network showed that, a suitable combination of the conditions for landslides occurrence represents siliciclastic deposits of Outer Western Carpathians (category 10 in Tab. 1) in the area of pastures, northwest oriented with slope from 9 to 15°. In Tab. 5 spatial distribution of susceptibility degree classes and slope deformations in them can be seen. From the table it is visible, that the neural network has identified more than 69 % of the area with very low and low degree of landslide hazard. Favorable conditions for the formation of slope deformation, with high and very high degree of landslide hazard pointed to less than 18% of the area.

For comparison, bivariate statistical analysis showed that the most susceptible ones are slope sediments, with a slope of 7° to 11° in the forest with slopes oriented to the south. Similarly, the multivariate analysis identified the slopes inclinations from 7° to 11 °, oriented to the north used as gardens as most vulnerable to slope deformation. From Tab. 6 it can be concluded, that the bivariate analysis shows spatial distribution of susceptibility degree classes very equally. Each class takes up about 20% of the area. In Tab. 7 multivariate analysis indicates more than 52% of the area with very low and low degree of landslide hazard. Favorable conditions for the formation of slope deformation, with high and very high degree of landslide hazard, are shown in more than 31 percent of the area.

Tab. 4 Spatial distribution of slope angle

Category	Interval [°]	Area [km²]	Area [%]
1	(< 2)	234.14	7.48
2	[2–3)	67.63	2.16
3	[3–5)	132.55	4.23
4	[5–7)	152.04	4.85
5	[7–11)	373.36	11.92
6	[11–17)	628.81	20.08
7	[17–20)	297.71	9.50
8	[20–31)	888.39	28.36
9	[> 31)	357.64	11.42

Tab. 5 Spatial distribution of landslide hazard degree and slope deformations based on neural network

Landslide hazard degree	Area [km²]	Area [%]	Area of landslides [km²]	Area of landslides [%]
1	1447.73	46.22	1.42	6.33
2	726.39	23.19	2.51	11.21
3	424.64	13.56	3.44	15.37
4	277.71	8.87	4.41	19.72
5	255.80	8.17	10.60	47.37
Σ	3132.26	100.00	22.38	100.00

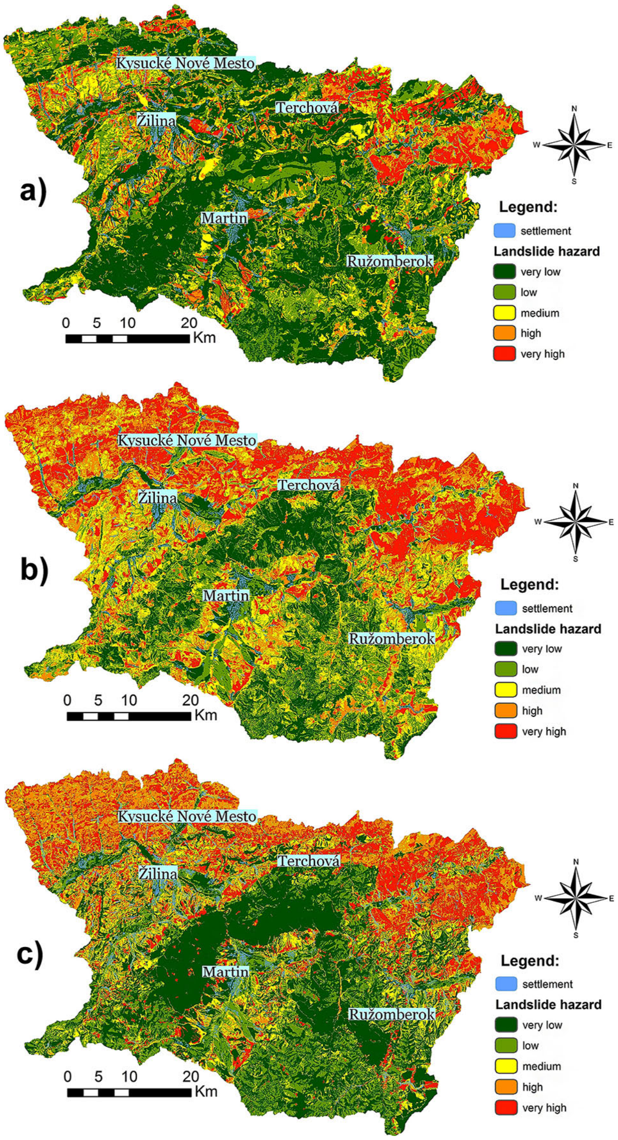

Fig. 4: Landslide hazard maps of Žilina region
a) using neural network; b) using bivariate statistical analysis; c) using multivariate statistical analysis

The ANN has many advantages compared to bivariate and multivariate statistical analysis. The ANN method is independent of the statistical distribution of the data and there is no need of specific statistical variables. Compared with the statistical methods, neural networks allow the target classes to be defined with much consideration to their distribution in the corresponding domain of each data source (Lee et al. 2004). The disadvantage of this method is the high technical and time complexity of the computer operations.

The resulting landslide hazard assessment has been prepared using prognostic raster maps with basic cell size of 10 to 10 m. Final dividing into liability classes took place an established "traffic light" system where 5 classes were distinguished: very low, low, medium, high and very high degree of landslide hazard. The resulting map constructed using ANN is shown in Fig. 4a. Map constructed using bivariate and multivariate analysis is shown in Fig. 4b and 4c.

Tab. 6: Spatial distribution of landslide hazard degree and slope deformations based on bivariate analysis

Landslide hazard degree	Area [km^2]	Area [%]	Area of landslides [km^2]	Area of landslides [%]
1	626.54	20.00	0.05	0.22
2	630.90	20.14	0.22	0.99
3	625.44	19.97	2.04	9.13
4	625.05	19.96	5.10	22.77
5	624.33	19.93	14.97	66.89
Σ	3132.26	100.00	22.38	100.00

Tab. 7: Spatial distribution of landslide hazard degree and slope deformations based on multivariate analysis

Landslide hazard degree	Area [km^2]	Area [%]	Area of landslides [km^2]	Area of landslides [%]
1	1138.96	36.36	0.00	0.00
2	498.25	15.91	0.07	0.31
3	493.50	15.76	0.95	4.25
4	506.70	16.18	4.45	19.86
5	494.86	15.80	16.91	75.57
Σ	3132.26	100.00	22.38	100.00

To verify the rate of success of created prognostic landslide susceptibility maps receiver operating characteristics (ROC) curves were used. The most important parameter is the area under curve (AUC). The size of the AUC determines the overall quality of predictive models. The maximum area of graph is 1 (ideal model, success rate is 100 %), the area of model with a success rate of 50 % has AUC = 0.5 (trivial model). The closer the area to the value 1 is, the more accurate the model is (Bednarik et al., 2010). Using bivariate statistical analysis the AUC is 0.852, for multivariate 0.919 and using NN the result is 0.924. The results shows that biggest rate of success has the prognostic landslide hazard map created using ANN and it is equal to 92.4 %. The calculated curves are shown in Fig. 5. According to the ROC curves only small differences between the results of used methods are visible. As follows from Fig. 4, there is a significant difference. Due to this, second approach of verification had been used based on overlaying prognostic models with the landslide inventory map.

This verification approach compared areas of registered landslides with high and very high degree of landslide hazard, it means with the classes 4 and 5 in the prognostic map. In this case, the results did not correspond to the results of ROC curves. The lowest success has a prognostic map created by neural network and is equal to 67.09 %. Prognostic map

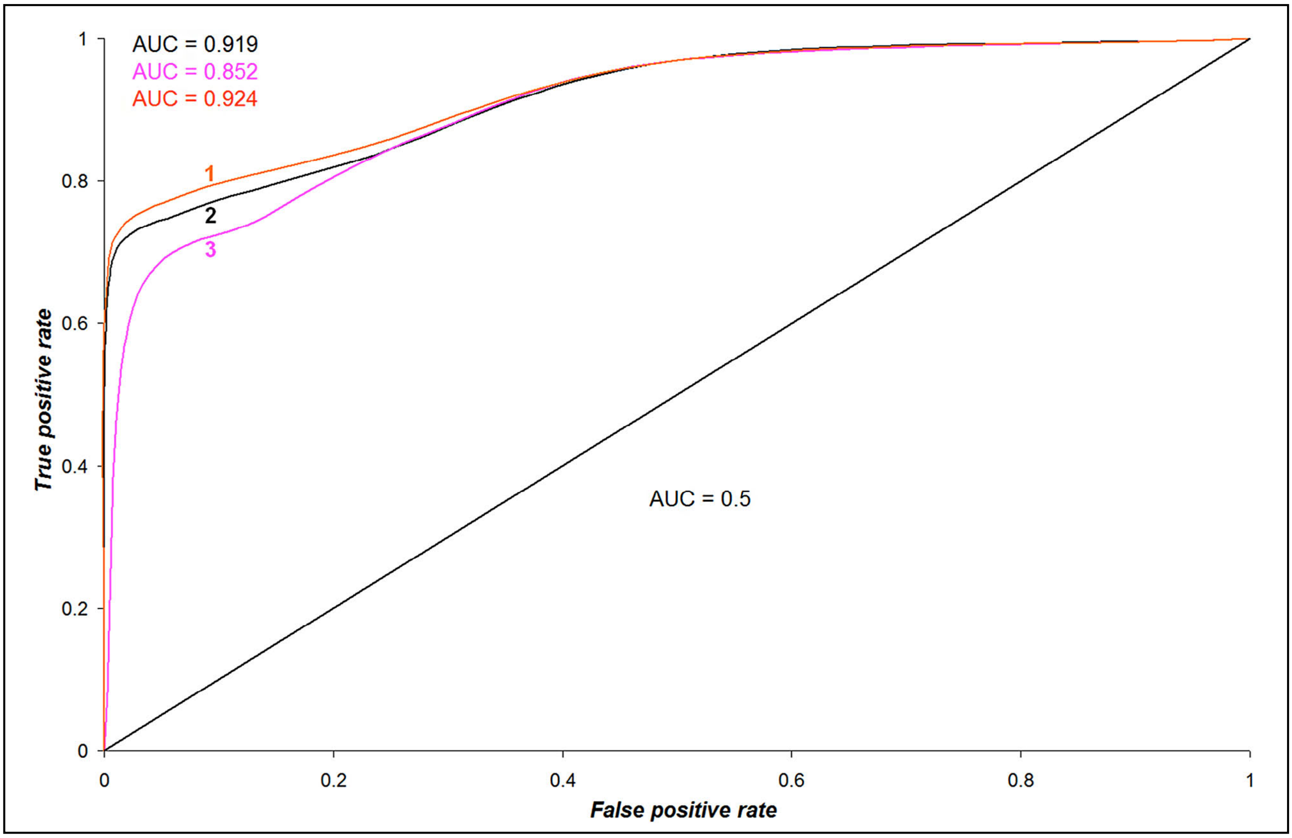

Fig. 5: ROC curves

1) artificial neural network; 2) multivariate statistical analysis; 3) bivariate statistical analysis

created using bivariate statistical analysis has a percentage success 89.65 %. The highest success was reached by prognostic map constructed using multivariate analysis and it is equal to 95.43 %. This result corresponds with figures much better than the result calculated by ROC curves. This can be caused by the fact that the neural network´s ROC curve was calculated in Matlab and the others were derived from data given by ArcGIS.

From the results we can conclude that bivariate statistical analysis shows the area as most unfavorable exactly like in the work of Tornyai & Dunčko, 2013. This result cannot be considered as incorrect, because the classification into landslide hazard classes was made by fully automated manner in ArcGIS environment, similarly as in the case of NN and multivariate statistical analysis in order to objectively compare the results. With manual adjustment, which, however, requires considerable experience the model could be "tuned in".

5. CONCLUSION

In the paper landslide hazard processing by the neural network was presented. Multilayer feed-forward neural network with back-propagation training algorithm was used. For training active landslides and stable areas were used. Preparation of input parameters and visualization of resulting maps were processed in ArcGIS, the actual calculation of the neural network was carried out in MATLAB. Statistical processing of landslide hazard assessment is based on the geological principle of actualism, which means that landslides will occur in places where they occurred in the past respectively in present under the similar activation conditions.

The resulting landslide hazard assessment has been prepared using prognostic raster maps with basic cell size of 10 to 10 m. Final dividing into liability classes adopted an established "traffic light" system.

Applied bivariate statistical analysis has found that the most favorable conditions for the development of slope deformation create the combination of Outer Western Carpathian slope sediments, south-oriented with slope angle from 11 to 17° in areas where land is used as a transitional woodland-shrub.

Applying multivariate conditional analysis few possible combinations of input parameters with a 100% probability of slope failures were identified. As an example a combination of Outer Western Carpathian slope sediments in the natural grasslands area with slope angle from 7 to 11° oriented to the North is provided.

Suitable combination of the conditions for landslides occurrence for artificial neural network represent siliciclastic deposits of Outer Western Carpathians in the area of pastures, northwest oriented with slope from 9° to 15°. To verify the rate of success of created prognostic landslide susceptibility maps receiver operating characteristics (ROC) curves were used.

Given that these methods (bivariate and multivariate) are well known and verified, the result of artificial neural network may be considered as good, however after comparing the results of the multivariate analysis showed the most accurate outcome. Better results can be achieved by manual editing of the final hazard classification, but this step requires considerable experience not only in the field of engineering geology.

Acknowledgements: This work was supported by the Slovak Research and Development Agency under the contract No. APVV-0129-12, APVV SK-CZ-2013-0141, VEGA agency no. 1/0131/14 and the Comenius University Grant UK/552/2015.

References

Aleotti P., Baldelli P., Polloni G. & Puma F., 1998: Keynote paper: Different approaches to landslide hazard assessment. *In:* Proc. Of the Second Conf. Environ management (ICEM – 2), Wollongong (Australia) 10–13 Feb. 1998, pp. 3–10.

Atkinson P.M. & Tatnall A.R.L., 1997: Introduction Neural networks in Remote Sensing. *International Journal of Remote Sensing*, 18, 4, 699–709.

Baum E. B. & Haussler D., 1989: What Size Net Gives Valid Generalization. *Neural Computation*, 1, 1, 151–160.

Bednarik M. & Pauditš P., 2010: Different ways of landslide geometry interpretation in a process of statistical landslide susceptibility and hazard assessment: Horná Súča (western Slovakia) case study. *Environmental Earth Sciences*, 61, 4, 733–739.

Bednarik M., Magulová B., Matys M. & Marschalko M., 2010: Landslide susceptibility assessment of the Kraľovany–Liptovsky Mikulaš railway case study. *Physics and Chemistry of the Earth*, 35, 162–171.

Dostál I., Putiška R. & Kušnirák D., 2014: Determination of shear surface of landslides using electrical resistivity tomography. *Contributions to Geophysics and Geodesy*, 44, 2, 133–147.

Ermini L., Catani F. & Casagli N., 2005: Artificial neural networks applied to landslide susceptibility assessment. *Geomorphology*, 66, 1-4, 327–343.

Grman D., Boszáková M., Magdošková M., Takáč P., Wanieková D., Žec B., Balážová R., Hajduková J., Syčev V., Udič M., Stercz M., Žurbej J., Bednarik M., Laho M., Liščák P., Páleník M., Sluka V., Smolka J., Smolka M., Šimeková J., Farkašovský R., Jacko S. & Janočko J., 2011: Inžinierskogeologické mapovanie svahových deformácií v najohrozenejších územiach flyšového pásma v M 1:10 000, Orientačný inžinierskogeologický prieskum, záverečná správa, Košice.

Hertz J. A., Krogh A. S. & Palmer A., 1991: Introduction to the Theory of Neural Computation. Addison Wesley Publishing Company, Redwood City, CA, 327 p.

Hrašna M., 1986: Mapy inžinierskogeologického rajónovania pre účely urbanizácie. *In:* Zborník referátov z vedeckej konferencie "Progresívne smery v inžinierskogeologickom výskume". Comenius University in Bratislava, Faculty of Natural Sciences, Department of Engineering Geology, Bratislava, pp. 68–82.

Lee S., Ryu J., Won J. & Park H., 2004: Determination and application of the weights for landslide susceptibility mapping using an artificial neural network. *Engineering Geology*, 71, 3-4, 289–302.

Liščák P., Pauditš P., Petro Ľ., Iglárová Ľ., Ondrejka P., Dananaj I., Brček M., Baráth I., Vlačiky M., Németh Z., Záhorová Ľ., Antalík M., Repčiak M. & Drotár D., 2010: Registration and evaluation of newly evolved slope failures in Prešov and Košice regions in 2010. *Mineralia Slovaca*, 42, 4, 393–406.

Macchi J. G. & Deravignone L., 2006: From GIS to Artificial Neural Networks: ten years studies on fortified villages in central Italy. *In*: Archäologie und Computer: Kulturelles Erbe und neue Technologien. Vienna, 7–10 November 2005, VIENNA, Phoibos Verlag, pp. 1–17.

Pauditš P., 2005: Hodnotenie náchylnosti územia na zosúvanie s využitím štatistických metód v prostredí GIS. PhD Thesis, Comenius University in Bratislava, 153 p.

Potfaj M. (Ed.), Šlepecký T., Maglay J., Hanzel V., Boorová D., Žecová K., Kohút M., Nagy A., Teťák F., Vass B., Sandanus M., Buček S., Sýkora M., Kohler E., Fejdiová O., Kandera K., Samuel O., Bubík M. & Beleš F., 2003: Vysvetlivky ku geologickej mape regiónu Kysuce 1: 50 000. ŠGÚDŠ, Bratislava, 193 p.

Pradhan B. & Lee S., 2009: Landslide risk analysis using artificial neural network model focussing on different training sites. *International Journal of Physical Sciences*, 4, 1, 1–15.

Putiška R., Dostál I., Kušnirák D. & Bednarik M., 2015: Landslide Geophysical Imaging - Kapušany Landslide Case Study. *In:* Menghini A. & Beard L. (Eds.): Near Surface Geoscience Turin, Italy, 6–10 September 2015, pp. 145–149.

Rumelhart D.E., Hinton G.E. & Williams R.J., 1986: Learning internal representations by error propagation. *In:* Rumelhart D.E. & McClelland J.L. (Eds.): Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1., Foundations. MIT Press / Bradford Books, Cambridge, MA, pp. 318–363.

Šilhán K., Růžek I., Burian L. & Putiška R., 2014: Landslide activity by the Kozárovce village (Levice district, Slovakia). *Geoscience Research Reports*, Volume 2014, 141–144.

Tornyai R. & Dunčko M., 2013: Použitie bivariačnej a multivariačnej analýzy na hodnotenie zosuvného hazardu Kysuckého regiónu. *Acta Geologica Slovaca*, 5, 2, 179–193.

White H., 1989: Learning in Artificial Neural Networks: A statistical Perspective. *Neural Computation*, 1, 4, 425–464.

Zhou W., 1999: Verification of the nonparametric characteristics of backpropagation neural networks for image classification. *IEEE Transaction on Geoscience and Remote Sensing*, 37, 2, 771–779.