

Research Article

Ivica Milevski, Slavoljub Dragičević, and Matija Zorn*

Statistical and expert-based landslide susceptibility modeling on a national scale applied to North Macedonia

<https://doi.org/10.1515/geo-2019-0059>

Received Feb 05, 2019; accepted Sep 20, 2019

Abstract: This article presents a Geographic Information System (GIS) assessment of Landslide Susceptibility Zonation (LSZ) in North Macedonia. Because of the weak landslide inventory, statistical method (frequency ratio) is combined with Analytical Hierarchy Process (AHP). In this study, lithology, slope, plan curvature, precipitations, land cover, distance from streams, and distance from roads were selected as precondition factors for landslide occurrence. There are two advantages of the approach used. The first is the possibility of comparing of the results and cross-validation between the statistical and expert based methods with an indication of the advantages and drawbacks of each of them. The second is the possibility of better weighting of precondition factors for landslide occurrence, which can be useful in cases of weak landslide inventory. The final result shows that in the case of weak landslide inventory, LSZ map created with the combination of both models provide better overall results than each model separately.

Keywords: landslides, LSZ, LSI, AHP, geo-environmental factors, North Macedonia

1 Introduction

The territory of the Republic of North Macedonia (hereafter North Macedonia) is highly exposed to various natural hazards, including landslides. Extensive areas are covered with erodible crystalline rocks (gneiss, mica schists, and other schists), sandstones, lacustrine and fluvial deposits.

For the topography (Figure 1) steep slopes are characteristic (the average slope is 15.5°). Climate is semi-arid (annual precipitation of 500–700 mm) and the vegetation is sparse. In addition to natural factors, human impact significantly contributes to the activation of landslides. Landslides are especially common on the valley sides, where Neogene lacustrine sands and sandstones are superimposed over inclined less permeable layers of clays and schists [1].

To reduce the risk from landslides, identifying, and mapping the landslides-prone area in the form of Landslide Susceptibility Zonation (LSZ) is very important [2–7]. In addition to the high frequency of landslides and yearly damage of up to several million euros, in North Macedonia only a few small-scale studies of LSZ were made [8–12]. Thus, this is the first attempt to prepare a LSZ map on a national scale. In addition, the results will help in the further development of the models and help speed up the compilation of a nationwide landslide inventory map.

Landslide researchers have elaborated many landslide susceptibility (LS) models [13, 14]. They are generally divided into qualitative or heuristic methods (partially subjective and essentially based on expert knowledge) and quantitative methods (based on numerical expressions of the relations between precondition (triggering) factors for landslide occurrence and landslide activity). In both the results are prone to uncertainties due to possible errors related to the pre-processing of triggering factors or model limitations [15]. Considering these the combination of both approaches in the form of hybrid methods is often used. These is also the case, when landslide dataset is not sufficient for a reliable statistical analysis [16]. Thus, the combined approach using data-driven Landslide Susceptibility Index (LSI) and an expert-based semi-quantitative (AHP) method is applied in this study. Instead of using experts' opinion only for the weighting of precondition factors for landslide occurrence in the AHP method, the bivariate analysis method (calculating the landslide density of each class) is additionally used to analyse the weights. In addition, relationship and dependence statistics are

*Corresponding Author: **Matija Zorn:** Research Centre of the Slovenian Academy of Sciences and Arts Ljubljana, Slovenia; Email: matija.zorn@zrc-sazu.si

Ivica Milevski: Cyril and Methodius University in Skopje, Faculty of Sciences, Institute of Geography, Skopje, North Macedonia

Slavoljub Dragičević: University of Belgrade, Faculty of Geography, Belgrade, Serbia

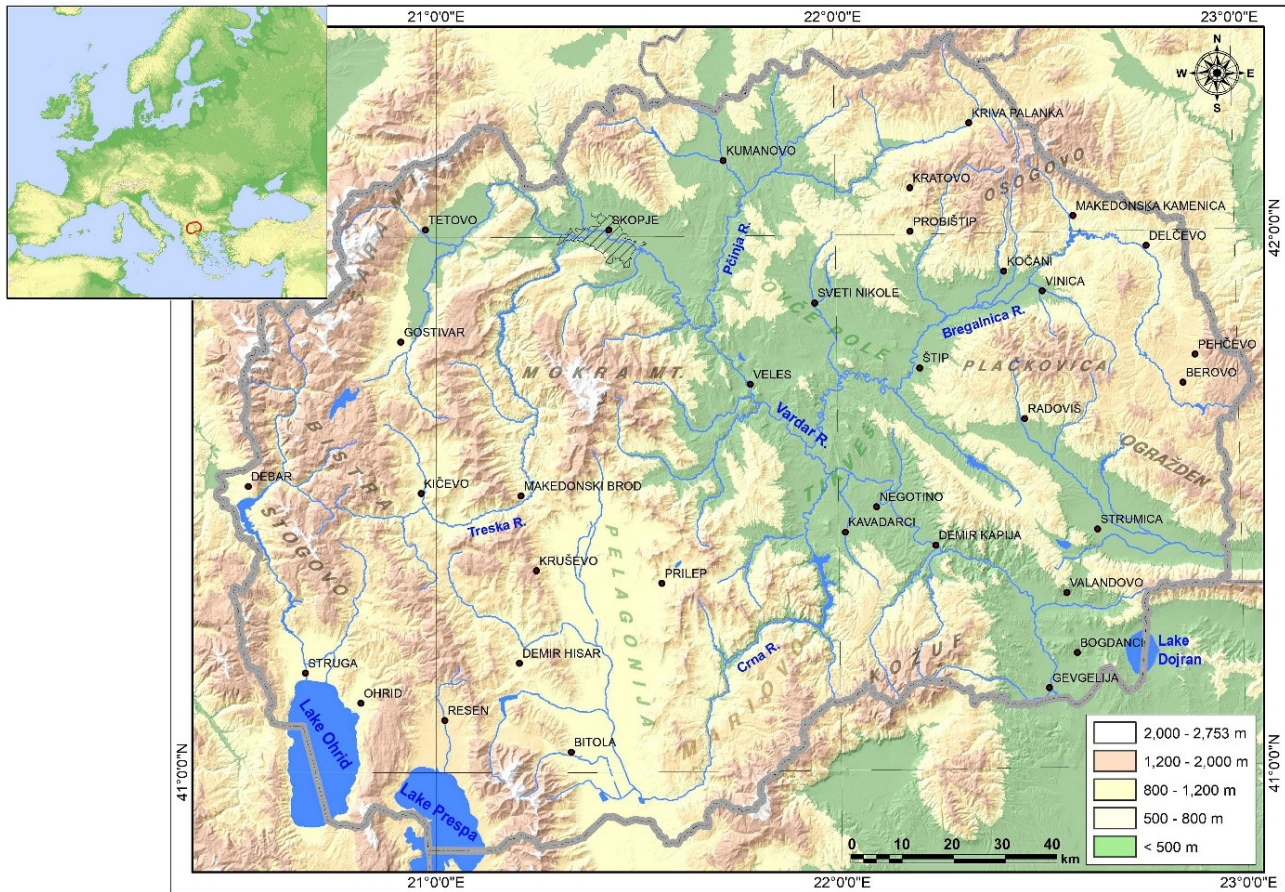


Figure 1: Geographic location and relief map of North Macedonia.

used to analyse the correlation coefficients of triggering factors.

2 Data and Methods

One of the basic requirements for accurate LS modelling is to have a sufficient spatial dataset (inventory) that represents former and recent landslides. This is the most critical information layer for the carrying out quantitative statistical analysis, and also for the validating of the model's accuracy [17, 18]. Unfortunately, in North Macedonia there is no such inventory yet (there are only partial records and data at some institutions). For this reason, we collected from various sources (e.g., geological maps, scientific papers and reports, media) as many landslide records as possible. The final landslide dataset consists of 302 landslides, which is rather a small number considering the study area. However, recently Chalkais *et al.* [18] prepared a GIS-based LS model of Peloponnese (in Greece) with a similar landslide density and acceptable accuracy (76%).

Most of the recorded landslides in our database are of the rotational and translational type, then rock-falls and earth-falls, and rarely the complex type. Some of the recorded landslides were new, and some were reactivated. Because of the small inventory, the landslide dataset was randomly split into two equal groups: a training dataset and a validation dataset. The training dataset was used to carry out the statistical analysis, and the validation dataset was used to verify the results.

Proper identification of precondition (triggering) factors for landslide occurrence is the basis for the Landslide Susceptibility Assessment (LSA). According to Crozier [3], depending on the characteristics of the study area, at least three factors must be included in GIS analysis, including topography, lithology, and land use. Donati and Turini [19] indicate that the most common precondition factors are: lithological units, tectonic features, slope angle, proximity to (road and drainage) networks, land cover, and rainfall distribution. At Landslide Susceptibility Mapping (LSM) at the national scale in Austria, the authors used six contributing factors [20]: elevation, slope angle, aspects, land cover, lithology, and Topographic Position

Index (TPI). The same number of factors (six) is used in the LSM at the national level in Slovenia [21] including: lithology, slope angle, surface curvature, land use, maximum 24-hour precipitation, and aspect. In a similar study at the national level in Slovenia [22] the following relevant precondition factors for landslide occurrence were identified: lithology, slope angle, land use, surface curvature, distance to structural elements, and aspect. In the regional GIS-based LSM of the Peloponnese, the authors selected seven factors as a most relevant: elevation, slope, aspect, mean annual precipitation, peak ground acceleration, lithology, and land cover [18].

In this study, eight precondition factors were taken into account considering the nature of the study area, the scale of the analysis, and data availability. These are topographic factors (slope, aspect, and plan curvature), lithology, precipitation, land cover, distance from streams, and distance from roads. The data for slope, aspect, and curvature were calculated from a 15 m Digital Elevation Model (DEM) of the entire country using SAGA GIS software. The 15 m DEM was prepared with filtering and resampling of the 5 m TIN-like DEM of the State Agency of Cadaster, resulting in a natural (non-triangular) appearance. Based on the distribution and structure of slopes in North Macedonia [1], slope values are classified into five classes ($0-5^\circ$, $5-10^\circ$, $10-30^\circ$, $30-45^\circ$, and more than 45°). In the analysis, four main aspects were considered (N, E, S, W), and the curvature is classified into five classes: highly convex, convex, flat, concave, and highly concave. The lithology map was prepared from a 1:100,000 digitalized and rasterized geological map of the country with ninety-three lithological units: from Precambrian gneiss and mica schist through Mesozoic limestone to Cenozoic sediments of marine, lacustrine, and riverine origin. These lithological units are generalized according to engineering-geological features and clustered into five classes, from clastic sediments to very resistant rocks (marble, limestone, quartzite, etc.). The land cover layer was prepared according to the CORINE (CLC2012) classification hierarchy (level 1). The precipitation map of the country (average sums for the period from 1986 to 2015) was prepared from the gauge data and vertical gradients [23]. Distances from the streams were derived using the topographic river network (25K) of the State Agency of Cadaster, and the distance from roads was prepared from a freely available Open Street Map (OSM) road network in vector (.shp) format. Following the methodology, five buffer zones for roads (with 50 m intervals) and streams (with 100 m intervals), which were considered most influential for landslide triggering [4–7], were created and rasterized (Figure 2). All the layers were then converted to raster grids with 15×15 m cells.

The next step was the selection of a suitable LSA method. One of the principal assumptions employed in the LSA is that the triggering factors for landslide occurrence in the future will stay the same as they were in the recent past. In general, LSA can be elaborated by means of qualitative and quantitative methods, or their combination [24]. Keeping in mind the large study area and a weak landslide inventory, a combination of frequency ratio (*Fr*) and Analytical Hierarchy Process (AHP) was used.

The frequency ratio method, which is among the best for larger areas [13–17], is based on the relationship between the spatial distribution of landslides and individual triggering factor [25]. This method calculates the Landslide Susceptibility Index (LSI) for each category *i* of all selected factors *j* (e.g., slope, lithology, land cover, etc.) selected for the study with the following equation [26]:

$$LSI_{ij} = \ln \left(\frac{N_{ij}}{A_{ij}} / \frac{N_T}{A_T} \right) \quad (1)$$

where N_{ij} is the number of landslides in category *j* of factor *i*, A_{ij} is the area of this category, N_T is the total number of landslides, and A_T is the total area under investigation. Thus, LSI represents the relative susceptibility to landslide occurrence. The final LSI map is calculated by summing up the values of all factors; that is by summing the values for each individual grid cell of all six digital layers. Next, we classified these continuous values into five classes (very low, low, medium, high, and very high LS zones). There is no widely accepted agreement concerning the best approach for classifying LS values. Natural breaks and quantile classification are the most frequently used [18, 24–26]. Both classifications were performed using SAGA GIS software, and their results are compared to ROC curve validation.

AHP is a semi-qualitative method that involves a matrix-based pairwise comparison of the weight (contribution) of various factors for landsliding. It was developed by Saaty [27] and has gained widespread attention [28]. The factor weight for each precondition factor is usually determined by an expert-based pairwise comparison matrix as described by Saaty and Vargas [29]. In our case, this matrix is compiled by the combination opinions of two experts, the field experience of the authors, and the results of the frequency ratio rankings. As with the LSI approach, in the AHP approach the final map is calculated by summing up the values of individual grid cell of all digital layers (see Figure 2). The same course of action is used for reclassifying the AHP values into different LS zones and map validation.

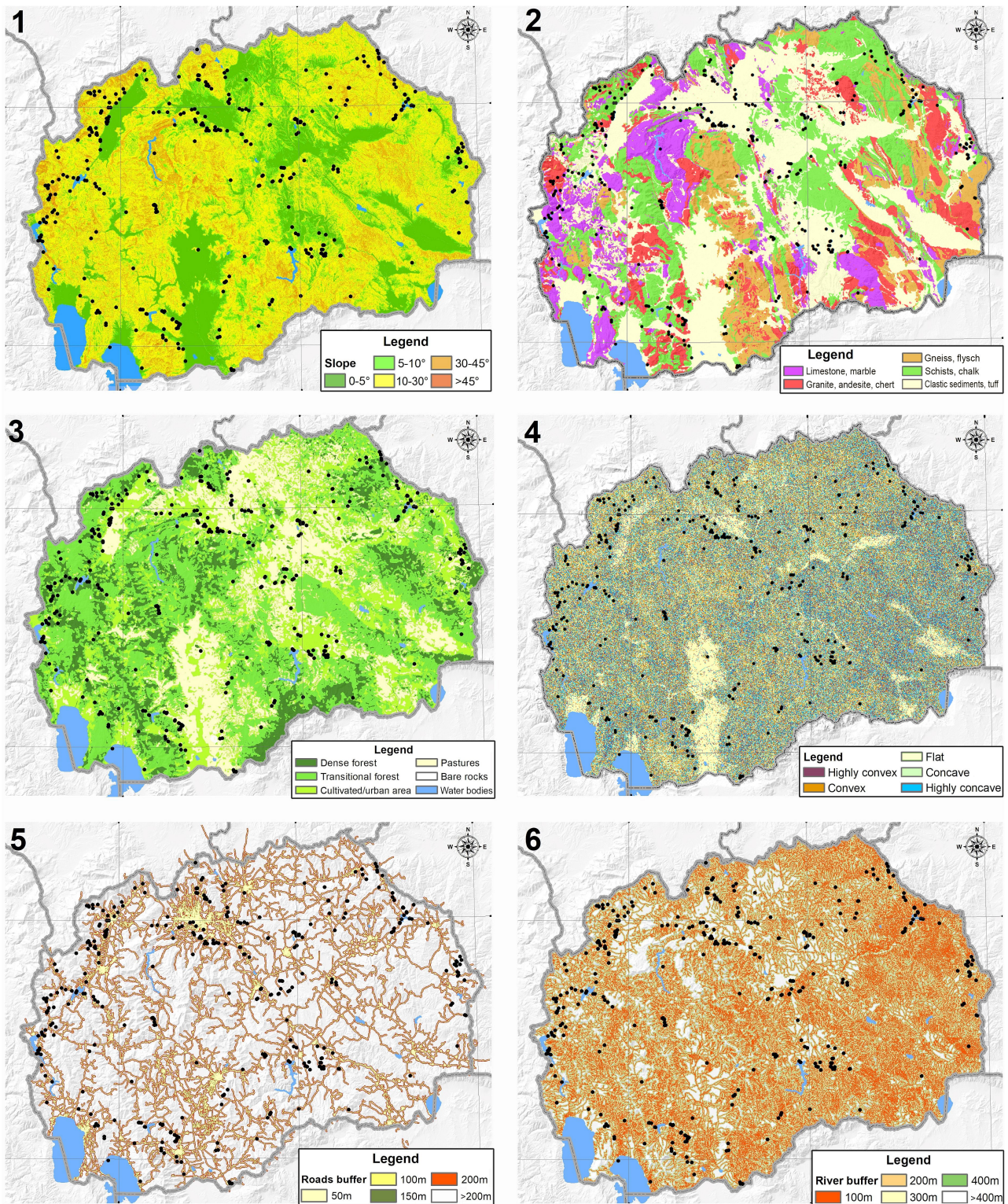


Figure 2: Raster layers of selected landslide-triggering factors: 1) slope, 2) lithology, 3) land cover, 4) planar curvature, 5) road network buffers, and 6) river network buffers. Black dots represent landslides from the landslide inventory.

3 Results

If an individual factor class correlates highly with the landslides in the dataset, then the area associated with this class will have a high positive LSI value (Table 1). A negative LSI value for a specific class is an indicator of low landslide density for this class [13, 20–22]. The results in Table 1 show that the highest positive LSI value has a 0 to 50 m road buffer. However, the reason for this may be the nature of the landslide dataset because a significant number of landslides and rockfalls in our inventory are on the roadside slopes. On the other hand, the relative ratio of significance (in the form of ranking in the last column) between the classes for individual factor (of layer) corresponds well to our expectations from field analysis.

However, for better evaluation of the importance of individual factor, the standard deviation, range, maximum value, and sums of negative and positive values were calculated (see Table 2). According to the values in Table 2, apart from the road factor (which has a medium range of extreme LSI values), the most important factors are slope, lithology, and land cover.

All factors except aspect have a value range greater than 0.8 and an LSI_{max} value greater than 0.25. In fact, aspects have a very low value range from the maximum positive to the maximum negative value of only 0.27. This means they have an almost insignificant influence on LS. For these reasons, aspect is excluded from further LSM. However, some studies disagree regarding the correlation between aspect and landslides [18, 30, 31]. With regard to precipitation, the corresponding LSI values show an inconsistent pattern, with the highest index for the class of 600 to 800 mm and a sharp drop for the class above 800 mm. We assume that this is related to a weak landslide inventory at the higher mountain elevations, where the mean annual precipitation exceeds 800 mm. Thus, there are only twenty-six landslides above 1,500 m in the inventory, but a much larger number is to be expected. Because of this uncertainty, precipitation is excluded from the frequency ratio modelling, at least until a better inventory (especially with elevation) is prepared.

Nevertheless, the ratios between the factors (except aspect) were used as indicators (together with the expert opinions) for the AHP method.

To obtain factor weights in the AHP, individual factor is rated against every other factor by assigning a relative dominant value between 1 and 9 to the intersecting cell. When the factor on the vertical axis is more important than the factor on the horizontal axis, this value varies between 1 (equally important) and 9 (extremely important, or domi-

nant). Conversely, the value varies between the reciprocals 1/2 (0.5) and 1/9 (0.11). Because we used seven parameters, the comparison matrix has forty-nine boxes (Table 3). The matrix-based weight of the factors, as well as the consistency ratio (CR) of the matrix, was calculated with the AHP Excel template [32].

Saaty [27] states that the CR must be less than 0.1 to accept the computed weights, otherwise the ratings should be re-evaluated. In Table 3 the CR is 0.02, indicating the acceptable consistency of the comparison matrix.

Furthermore, the weight of each factor is multiplied by its rankings R (based on the previous LSI values of frequency ratio and expert rankings [33]). The results are presented in Table 4.

The final maps of quantitative LSI and semi-qualitative AHP approaches are presented in Figures 3 and 4. At first glance they are similar, but some considerable differences are also present. Namely, in the LSI map (Figure 3), very high susceptibility areas are mostly near the roads in hazardous terrains (steep slope, unstable rocks, and weak vegetation). This is because of the high LSI value for the 0 to 50 m road buffer class is calculated from the training landslide dataset, which also influences the resulting map.

On the other hand, the AHP-based map (Figure 4) has a more visible dominance of the high susceptibility zone with increasing slope, rock erodibility, weak vegetation, and higher precipitations.

Because of the advantages of both approaches and assuming that some of the disadvantages of the both can be avoided, the next step overlays both raster grids. With this procedure, the mean sum of LSI and AHP values for individual cell was calculated $((LSI+AHP)/2)$ and the output grid was classified for quantiles and natural breaks. The final (combined) map is presented in Figure 5.

In order to choose the final LSZ map, a cross-validation technique was used to compare known landslide location data and the LSZ map. This is produced with analytical tables and the Receiver Operating Characteristics (ROC) curve.

In the case of LSI method both classifications (quantiles and natural breaks) show good accuracy of the implemented frequency ratio model because more than 60% of the landslides in the validation subset are in the classes of high and very high LS. However, quantile classification is somewhat better as only nine (5.9%) landslides fall into the classes of low and very low susceptibility, versus thirty (19.2%) landslides in the natural breaks classification (Table 5). Moreover, as many as 122 landslides (80.8%) fall into the classes of high and very high susceptibility, versus only ninety-five landslides (62.3%) in the natural breaks

Table 1: Landslide number (n), landslide density (in %), area, LSI, and ranking (5–1) for individual factor class.

| Factor/class | Landslide <i>n</i> | Landslide % | Area km ² | Area % | LSI | Rank |
|-------------------------------|--------------------|-------------|----------------------|--------|-------|------|
| <i>Slope</i> | | | | | | |
| 0–5 | 5 | 3.3 | 5,831.5 | 22.9 | –1.93 | 1 |
| 5–10 | 13 | 8.6 | 3,517.4 | 13.8 | –0.47 | 2 |
| 10–30 | 110 | 72.8 | 12,689.7 | 49.9 | 0.38 | 5 |
| 30–45 | 21 | 13.9 | 3,136.6 | 12.3 | 0.12 | 3 |
| > 45 | 2 | 1.3 | 274.7 | 1.1 | 0.20 | 4 |
| <i>Lithology</i> | | | | | | |
| Limestone, marble | 3 | 2.0 | 3,002.0 | 11.8 | –1.78 | 1 |
| Granite, andesite, chert | 5 | 3.3 | 1,334.0 | 5.2 | –0.46 | 2 |
| Gneiss, flysch | 32 | 21.2 | 5,800.3 | 22.8 | –0.07 | 3 |
| Schists, chalk | 42 | 27.8 | 7,280.9 | 28.6 | –0.03 | 4 |
| Clastic sediments, tuff | 69 | 45.7 | 8,032.7 | 31.6 | 0.37 | 5 |
| <i>Land cover</i> | | | | | | |
| Water bodies | 0 | 0.0 | 484.4 | 1.9 | / | |
| Dense forest | 1 | 0.7 | 591.9 | 2.3 | –1.26 | 2 |
| Transitional forest | 34 | 22.5 | 7,491.8 | 29.4 | –0.27 | 3 |
| Cultivated/urban area | 25 | 16.6 | 5,198.5 | 20.4 | –0.21 | 4 |
| Pastures, bare rock | 91 | 60.3 | 11,683.4 | 45.9 | 0.27 | 5 |
| <i>Curvature</i> | | | | | | |
| Highly convex | 21 | 13.9 | 5,530.3 | 21.7 | –0.45 | 1 |
| Convex | 32 | 21.2 | 5,804.8 | 22.8 | –0.07 | 2 |
| Flat | 28 | 18.5 | 4,688.7 | 18.4 | 0.01 | 3 |
| Concave | 42 | 27.8 | 4,895.4 | 19.2 | 0.37 | 5 |
| Highly concave | 28 | 18.5 | 4,530.8 | 17.8 | 0.04 | 4 |
| <i>Stream (buffer)</i> | | | | | | |
| 0–100 m | 74 | 49.0 | 8,829.4 | 34.7 | 0.35 | 5 |
| 100–200 m | 41 | 27.2 | 6,709.0 | 26.4 | 0.03 | 4 |
| 200–300 m | 15 | 9.9 | 3,570.3 | 14.0 | –0.35 | 2 |
| 300–400 m | 11 | 7.3 | 2,293.6 | 9.0 | –0.21 | 3 |
| > 400 m | 10 | 6.6 | 4,047.7 | 15.9 | –0.88 | 1 |
| <i>Road (buffer)</i> | | | | | | |
| 0–50 m | 44 | 29.1 | 1,801.3 | 7.1 | 1.42 | 5 |
| 50–100 m | 14 | 9.3 | 1,084.7 | 4.3 | 0.78 | 4 |
| 100–150 m | 16 | 10.6 | 1,313.6 | 5.2 | 0.72 | 3 |
| 150–200 m | 8 | 5.3 | 1,812.3 | 7.1 | –0.30 | 2 |
| > 200 m | 69 | 45.7 | 19,438.2 | 76.4 | –0.51 | 1 |
| <i>Precipitation</i> | | | | | | |
| 400–500 mm | 6 | 4.0 | 941.7 | 3.7 | 0.08 | 2 |
| 500–600 mm | 35 | 23.2 | 5,140.9 | 20.2 | 0.14 | 3 |
| 600–700 mm | 40 | 26.5 | 5,471.8 | 21.5 | 0.21 | 4 |
| 700–800 mm | 47 | 31.1 | 5,980.8 | 23.5 | 0.28 | 5 |
| 800–1200 mm | 23 | 15.2 | 7,940.4 | 31.2 | –0.72 | 1 |
| <i>Aspect</i> | | | | | | |
| N (0–45°; 315–360°) | 29 | 19.0 | 5,599.0 | 22.0 | –0.15 | 3 |
| E (45–135°) | 46 | 30.1 | 6,947.9 | 27.3 | 0.10 | 4 |
| S (135–225°) | 42 | 27.5 | 6,260.7 | 24.6 | 0.11 | 5 |
| W (225–315°) | 34 | 22.2 | 6,642.5 | 26.1 | –0.16 | 2 |

Table 2: Frequency ratio values (LSI) for individual factor showing its importance as a precondition for landslide occurrence.

| Factor | LSI _{min} | LSI _{max} | LSI _{range} | LSI _{stdv} | LSI _{neg.val} | LSI _{pos.val} |
|----------------|--------------------|--------------------|----------------------|---------------------|------------------------|------------------------|
| Slope | -1.93 | 0.38 | 2.31 | 0.95 | -2.41 | 0.70 |
| Lithology | -1.78 | 0.37 | 2.15 | 0.83 | -2.34 | 0.37 |
| Land cover | -1.26 | 0.27 | 1.53 | 0.64 | -1.74 | 0.27 |
| Road buffer | -0.51 | 1.42 | 1.93 | 0.80 | -0.81 | 2.91 |
| Stream buffer | -0.88 | 0.35 | 1.22 | 0.45 | -1.43 | 0.37 |
| Precipitations | -0.72 | 0.28 | 1.00 | 0.42 | -0.72 | 0.71 |
| Curvature | -0.45 | 0.37 | 0.82 | 0.29 | -0.52 | 0.42 |
| Aspect | -0.16 | 0.11 | 0.27 | 0.13 | -0.31 | 0.21 |

Table 3: AHP comparison matrix for selected factors.

| Factor | Slope | Lithology | Land cover | Precipitation | Roads | Curvature | Streams | Weight |
|---------------|-------|-----------|------------|---------------|-------|-----------|---------|--------|
| Slope | 1 | 2 | 3 | 4 | 5 | 5 | 6 | 0.349 |
| Lithology | 0.50 | 1 | 2 | 3 | 5 | 5 | 5 | 0.248 |
| Land cover | 0.33 | 1 | 1 | 2 | 3 | 3 | 4 | 0.152 |
| Precipitation | 0.25 | 0.33 | 0.50 | 1 | 2 | 3 | 3 | 0.103 |
| Roads | 0.20 | 0.20 | 0.33 | 0.50 | 1 | 1 | 2 | 0.058 |
| Curvature | 0.20 | 0.20 | 0.33 | 0.33 | 1 | 1 | 1 | 0.049 |
| Streams | 0.17 | 0.20 | 0.25 | 0.33 | 0.50 | 1 | 1 | 0.042 |

Table 4: Weight values of factors used for the AHP model.

| Factor | Value | Factor | Value |
|-----------------------------|-------|--------------------------------|-------|
| Slopes w = 0.349 | | Precipitation w = 0.103 | |
| 0–5° | 0.35 | 400–500 mm | 0.10 |
| 5–10° | 0.70 | 500–600 mm | 0.21 |
| 10–30° | 1.75 | 600–700 mm | 0.31 |
| 30–45° | 1.05 | 700–800 mm | 0.41 |
| > 45° | 1.40 | 800–1200 mm | 0.52 |
| Lithology w = 0.248 | | Roads w = 0.058 | |
| Clastic sediments, tuff | 1.24 | 0–50 m | 0.29 |
| Schists, chalk | 0.99 | > 50 m | 0 |
| Gneiss, flysch | 0.74 | Curvature w = 0.049 | |
| Granite, andesite, chert | 0.50 | Concave | 0.25 |
| Limestone, marble | 0.25 | Highly concave | 0.20 |
| Land cover w = 0.152 | | Flat | 0.15 |
| Pastures, bare rock | 0.61 | Convex | 0.10 |
| Cultivated/urban area | 0.46 | Highly convex | 0.05 |
| Transitional forest | 0.30 | Streams w = 0.042 | |
| Dense forest | 0.15 | 0–100 m | 0.21 |
| Water bodies | 0 | > 100 m | 0 |

Table 5: Data for LSI and AHP-based LSZ according to the quantile and natural breaks classification.

| LS class | LSI | | | | AHP | | | | LSI+AHP | | | |
|--------------|--|-------------|--------------|-------|---------------------------------------|-------------|--------------|-------|--------------|-------------|--------------|-------|
| | Quantile classification | | | | Natural breaks (Jenks) classification | | | | | | | |
| | Area % | Landslide n | % | Ratio | Area % | Landslide n | % | Ratio | Area % | Landslide n | % | Ratio |
| very low | 20.3 | 4 | 2.6 | 0.13 | 20.0 | 6 | 4.0 | 0.20 | 13.2 | 2 | 1.3 | 0.10 |
| low | 20.2 | 5 | 3.3 | 0.16 | 20.0 | 12 | 7.9 | 0.40 | 20.5 | 4 | 2.6 | 0.13 |
| moderate | 20.0 | 22 | 14.6 | 0.73 | 20.0 | 24 | 15.9 | 0.80 | 19.8 | 15 | 9.9 | 0.50 |
| high | 19.8 | 52 | 34.4 | 1.74 | 20.1 | 45 | 29.8 | 1.48 | 24.1 | 39 | 25.8 | 1.07 |
| very high | 19.7 | 70 | 46.4 | 2.35 | 20.0 | 64 | 42.4 | 2.12 | 22.4 | 91 | 60.3 | 2.69 |
| Total | 100.0 | 151 | 100.0 | | 100.0 | 151 | 100.0 | | 100.0 | 151 | 100.0 | |
| | Natural breaks (Jenks) classification | | | | | | | | | | | |
| very low | 40.9 | 8 | 5.3 | 0.13 | 20.0 | 5 | 3.3 | 0.17 | 22.8 | 4 | 2.6 | 0.11 |
| low | 19.8 | 22 | 13.9 | 0.70 | 13.3 | 12 | 7.9 | 0.59 | 17.9 | 6 | 4.0 | 0.22 |
| moderate | 20.4 | 28 | 18.5 | 0.91 | 26.6 | 24 | 15.9 | 0.60 | 26.0 | 20 | 13.2 | 0.51 |
| high | 11.7 | 63 | 41.1 | 3.51 | 30.0 | 67 | 44.4 | 1.48 | 14.9 | 37 | 24.5 | 1.64 |
| very high | 7.2 | 32 | 21.2 | 2.93 | 10.1 | 43 | 28.5 | 2.82 | 18.5 | 84 | 55.6 | 3.01 |
| Total | 100.0 | 151 | 100.0 | | 100.0 | 151 | 100.0 | | 100.0 | 151 | 100.0 | |

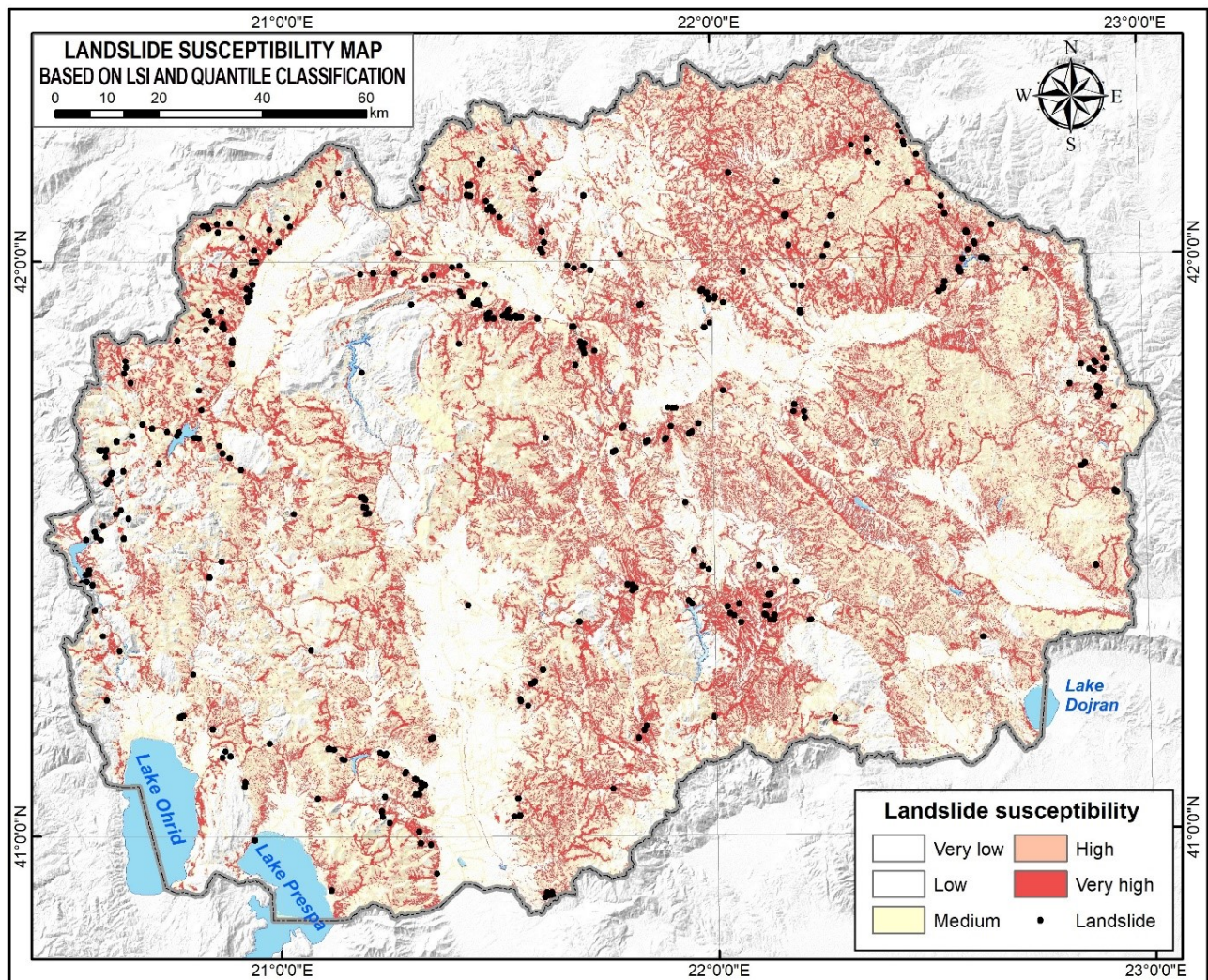


Figure 3: LSI map of North Macedonia (quantile classification).

classification, covering 39.5% and 18.9% of the total area of North Macedonia respectively. However, the number of landslides (in %) compared with the area (in %) of the high and very high susceptibility classes show a somewhat better ratio for natural breaks, versus quantile classification (*i.e.*, 6.4 and 4.1 respectively).

The analyses for the AHP approach show a somewhat different situation. Thus, according to the quantile classification of the AHP model, 72.2% or 109 out of 151 landslides in the validation subset, belongs to the high and very high LS classes covering 40.1% of the total area of North Macedonia (Table 5). The natural breaks classification shows the same, with 72.9% of the landslides located in the high and very high LS classes, which in turn cover 40.1% of the country. Moreover, the number of landslides (in %) compared with the area (in %) of the high and very high susceptibility classes show a somewhat better ratio for natu-

ral breaks, versus quantile classification (*i.e.*, 4.3 and 3.6 respectively). On the other hand, the same ratio for the areas with low and very low susceptibility classes is 0.76 and 0.6, respectively, which favors the quantile classification (lower value). Given that the ratio tends to be as close to 0 for the very low LS class, and that it increases well above 1 for the very high LS class, then natural breaks classification shows slightly better overall accuracy. Thus, according to the AHP map prepared with the natural breaks classification, 88.8% of the validated landslide fall within the very high, high, and moderate LS zones.

The third approach as a combination (averaging) of LSI and AHP sums of values retains some advantages of both models, which means that areas with high to very high and low to very low LS rates tend to be more distinctive (depending on the classification). Thus, with the implemented natural breaks scheme, there are 121 landslides

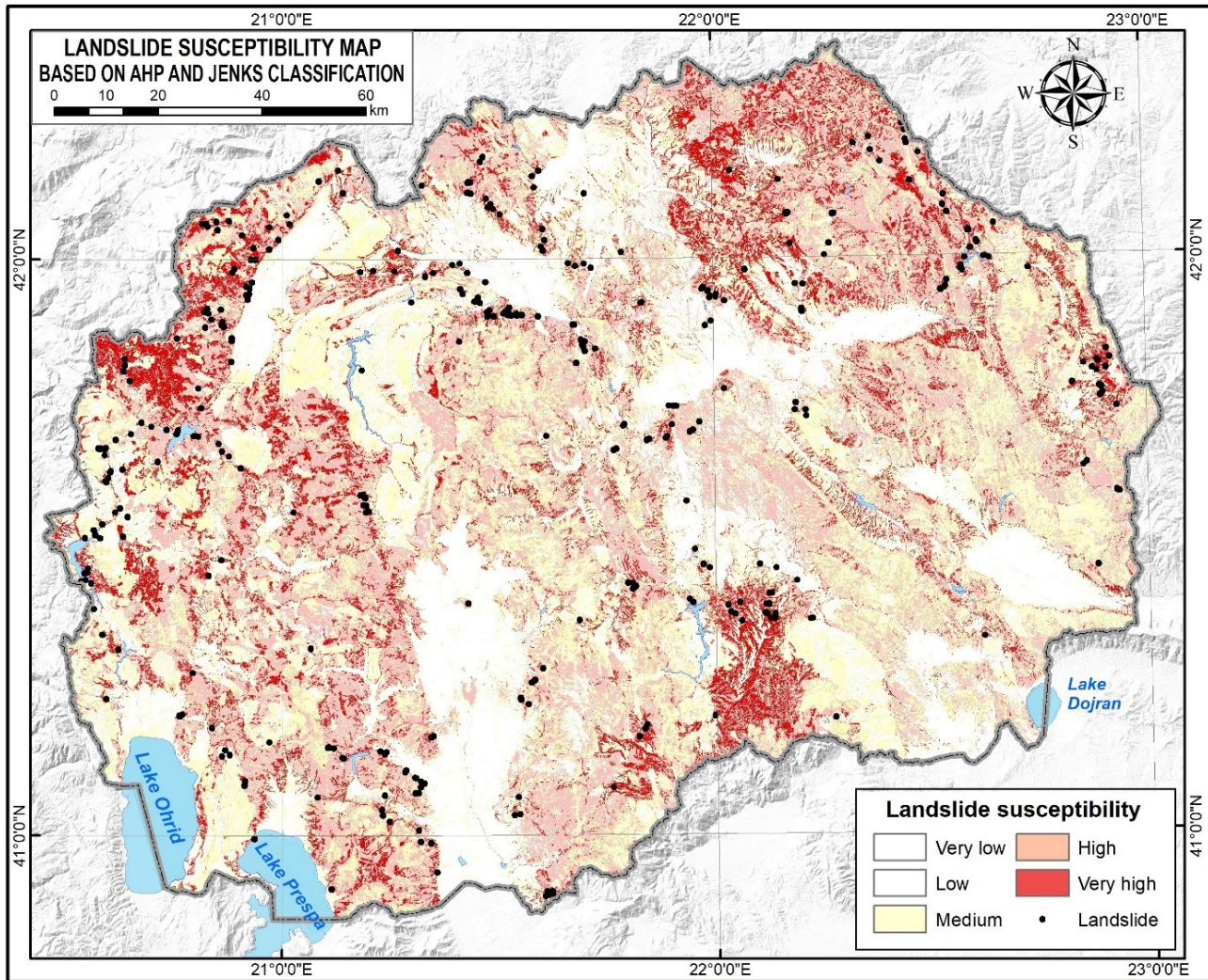


Figure 4: AHP-based landslide susceptibility map of North Macedonia (natural breaks classification).

(80.2%) in the classes of very high and high LS, which in turn cover only 33.4% of the country's area. Because of this, the ratio of the number of landslides (in %) and the area of class (in %) is 3.01 for the very high LS zone, which is the best result, whereas together with the high susceptibility zone it is 4.65 (only LSI is slightly better, at 6.44). Moreover, the ratio for the very low and low susceptibility zones is only 0.11 and 0.22 (or 0.33 in total), which is among the lowest values and indicates a small number of landslides in these zones (only 10 or 6.6%) compared to the total area (40.7%) (Table 5).

In order to estimate the overall performance of LS models in the study area, validation of the susceptibility maps was checked using the ROC curve, and the Area Under Curve (AUC) [34]. For better assessment of the model, except the spatially random half of the recorded (151) true-positive landslides (value 1) in the validation dataset, 450

false-positive landslides (value 0) were carefully selected. The ROC curve and AUC in this study were calculated using SPSS-statistical software and are presented in Figure 6. It is interesting that all three models (the LSI, AHP, and combined LSI+AHP-based model) have almost the same AUC with a very slight advantage of LSI with AUC at 0.790 vs. 0.709 vs. 0.784, respectively.

However, when the classification method is examined, natural breaks perform best in the combined LSI+AHP model (0.784), after which are LSI (0.781) and AHP (0.767). In the case of quantile classification, LSI performs best (0.783) compared to the AHP model and combined (LSI+AHP) model (0.696 vs. 0.766). All methods (LSI, AHP, and LSI+AHP) show nearly 78% agreement with landslide locations of the validation dataset, which is a reasonable result on this scale [16–18].

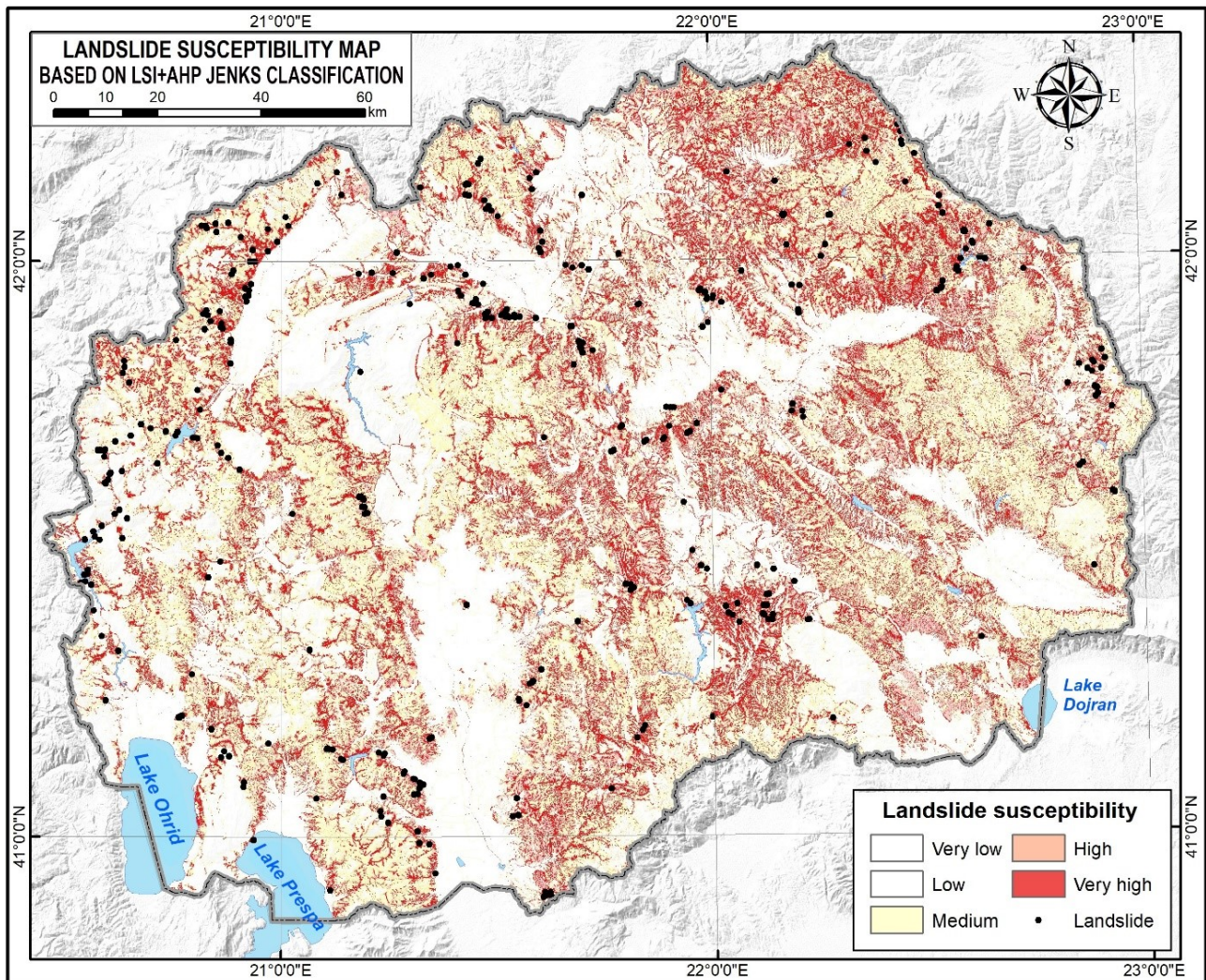


Figure 5: Combination of LSI (frequency ratio) and AHP-based landslide susceptibility map of North Macedonia (natural breaks classification).

Based on both LSZ models, most of the landslides (in the test and validation dataset) in North Macedonia occur on moderate slopes ($10\text{--}30^\circ$) and on terrain composed of clastic sediments (Neogene lacustrine deposits and colluvium sediments) and schists (mica schists, green schists, etc.). In addition, a significant number of landslides occur on terrains with weak vegetation (pastures, grasslands, and bare and erodible rocks), and also on the cultivated land on steep terrains and in urban areas. A statistically substantial number of landslides are located at a distance up to 100 to 200 m from the streams and up to 50 m from the roads, mostly as a large roadside rockfalls. Regionally, most of the area with high and very high LS in North Macedonia extends over hilly terrain and mountain foothills, on the sides of valley bottoms in gorges, and on the sides of depressions and basins that are usually covered by Neogene

lacustrine sediments (see Figures 3–5). Thus, according to the maps, the areas in the central part of the country (the Tikveš Depression), the north-east part of the hillslopes of Mount Osogovo and Mount Bilino, and the upper Bregalnica catchment (Figure 7), and the foothills of Mount Šara are among the most susceptible to landslides.

In contrast, larger plains and terrains composed of solid rock (limestone, marble, andesite, etc.), especially in the western part, shows low LS. However, field studies have revealed that even the occurrence of (smaller) landslides is not totally excluded near channels, roads, constructions, and other sites with substantial anthropogenic activities [1, 33].

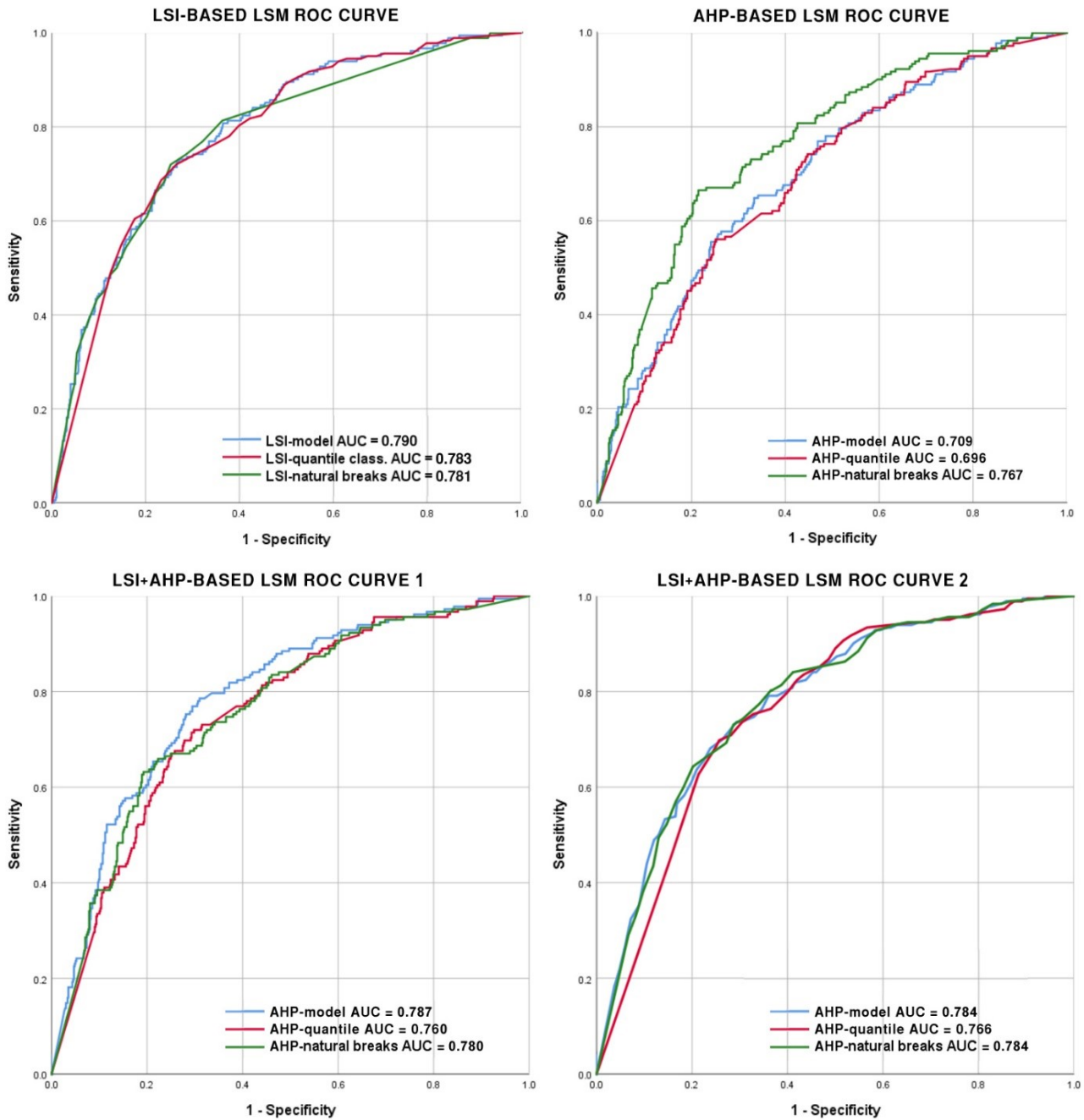


Figure 6: ROC curve and AUC for the LSI (upper left), AHP (upper right), and combined LSI+AHP-based LS maps: with (lower left) and without precipitation as a precondition factor (lower right).

4 Discussion

Large scale mapping of landslides-prone areas is based on the data obtained by extensive and expensive geotechnical research. Such detailed maps contain all the information needed for spatial planning. However, such detailed geotechnical investigations are not feasible (due to

expenses and time needed) for the country scale. Thus, for this scale an alternative and less time consuming solution for LSZ is needed [13–18, 35–37].

Many models for LSZ were elaborated using e.g., multivariate statistics, binary statistics, the AHP, neural networks, etc. All have some advantages and disadvantages in regards to the reliability of data sources, pre-processing of precondition factors, weighting of precondition factors,

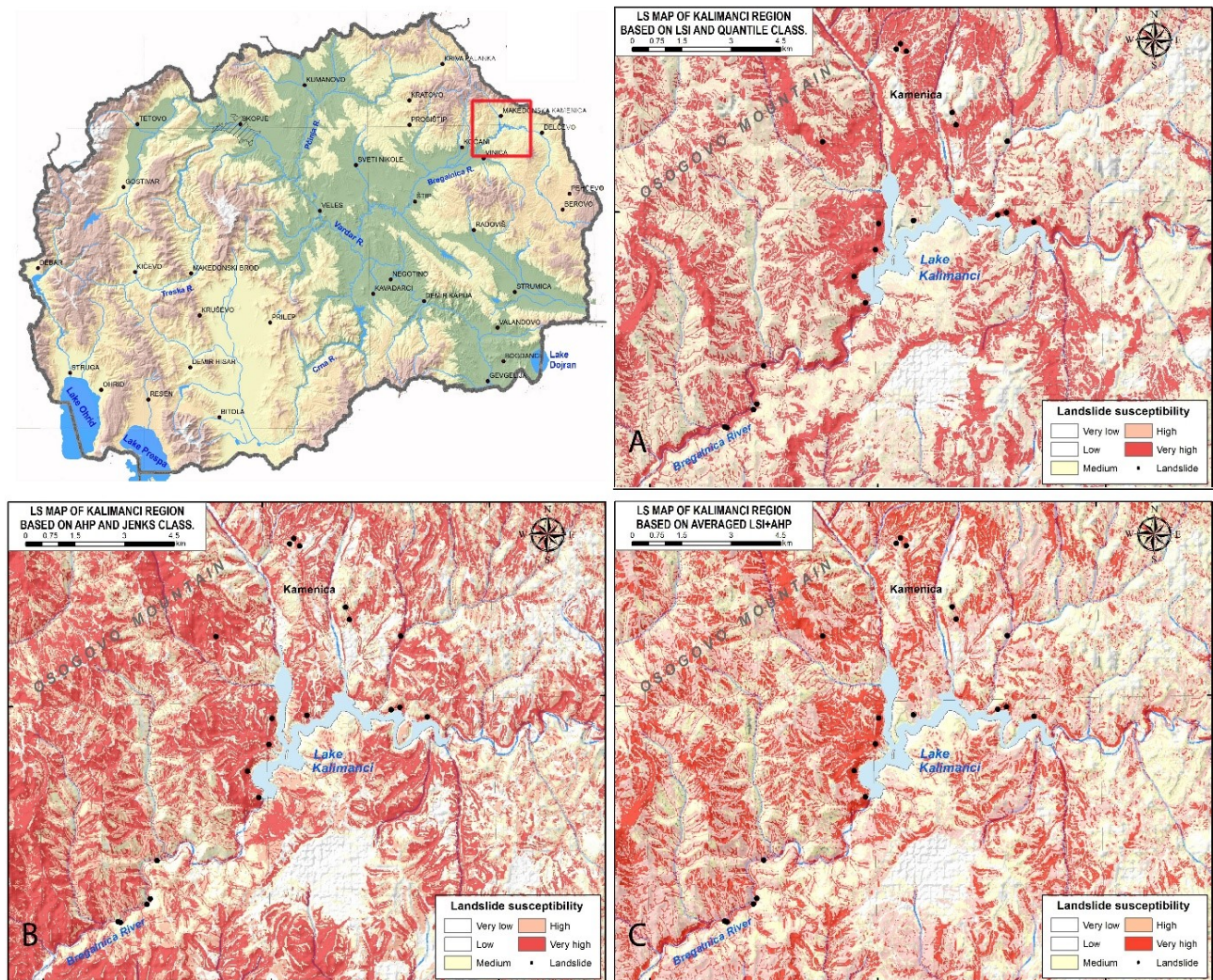


Figure 7: Part of the: A) LSI, B) AHP, and C) combined LSI+AHP LS maps in the Lake Kalimanci area (indicated with a red square on the country map; upper left).

level of subjectivity, etc. The article focuses on two such approaches, *i.e.*, on a statistical (LSI) method and an expert-based (AHP) method. Not to use experts' opinions to build a weighting matrix for precondition factors in the AHP, a bivariate analysis method (considering the landslide density of each class) was used to analyze the weights. Furthermore, relationship and dependence statistics were used to analyze the correlation coefficients of precondition factors.

For this study, eight precondition factors for landslide occurrence (slope, lithology, land cover, precipitation, plan curvature, aspect, distance from streams, and distance from roads) were selected, rasterized, and harmonized to a cell size of $15\text{ m} \times 15\text{ m}$. Then, the frequency ratio of landslide events in individual factor class was calculated in form of LSI values. These values were considered as the basis for the relative rankings of factor classes (con-

firmed or slightly corrected with expert rankings), which was then used in the AHP weighting matrix. However, because of the weak frequency ratio values, aspect is excluded from the modelling, whereas precipitation is considered only in the AHP approach (because of inconsistent ratio with precipitation increase). Thus, the final LSI and AHP models are the sums of grid cell values for each of the remaining six factors or layers. With further quantile and natural breaks classification, five classes of LSZ are defined and represented in the form of a LS map. Even with a very limited landslide inventory, statistically, there is about 78% agreement (AUC value) between prepared LS maps and 151 landslide locations (true positive) from the validation dataset. It is interesting that both approaches, LSI and AHP, show a very similar AUC in slight favor of LSI and quantile classification. Combined together (with values averaged), very consistent results are produced, at

least with respect to our previous field research [1, 8–10] and knowledge. These results show that a significant area of the country (18.9% by the LSI model, 33.4% by the LSI+AHP model, and 40.1% by the AHP model) falls into high and very high LS zones, which are verified with the landslide dataset.

The approach used has several limitations. As the modelling is based on national-scale datasets, the results are unsuitable for a reliable site-oriented spatial analysis, for which a much more detailed landslide inventory as well as precondition factors in better resolution are needed. *E.g.*, for this study the geology data are from the geological (lithological) map with a spatial resolution of approximately 80 to 100 m and the DEM used has a spatial resolution of 15 m. Furthermore, the resulting LS maps present only the predicted spatial distribution of landslides and not also the temporal probability of landslide occurrence (for which the exact data about the precipitation patterns and highest intensities must be considered). Further limitation is a very complex relationships between the precondition (triggering) factors. Thus, approach used can only be applied as a first step of LS assessment (accurate with mentioned limitations only for national scale). The approach is simple and makes a meaningful use of limited datasets, as shown in some other studies [38].

To improve the approach some points must be considered, *e.g.* a significant improvement of the landslide data inventory and a spatial resolution of the precondition data layers, an improvement of the factors weighting, the inclusion of additional precondition factors (*e.g.*, Topographic Wetness Index (TWI), Stream Power Index (SPI), Normal Difference Vegetation Index (NDVI), geological structural elements, precipitation intensities, and the introduction of alternative model validation approaches [36–38].

5 Conclusion

The goal of LS maps is not merely to show hazardous areas, but to help in activities to reduce hazard. If applied properly, such maps can help to avoid future damage. This is especially important when climate change and its consequences in terms of higher landslide activity are considered. In North Macedonia, as in some other European countries [39, 40], national funds are primarily used for the recovery of landslide damage, and much less for the prevention (also for the elaboration of LS maps). In this sense, approach used is the first attempt in North Macedonia at the country level, and it is hoped that further improvements will be made soon. Overall, this approach

can produce reliable LS maps up to the regional scale and provide useful information for the decision making in regional planning.

References

- [1] Milevski I., General geomorphological characteristics of the Republic of Macedonia. *Geographical Reviews*, 2015, 48, 5-25
- [2] Varnes D.J., *Landslide Hazard Zonation: A Review of Principles and Practice*. United Nations Educational, Scientific and Cultural Organization, Paris, 1984
- [3] Crozier M. *Landslides: Causes, Consequences and Environment*. Croom Helm, London, 1986
- [4] Soeters R., van Westen C.J., Slope instability recognition, analysis, and zonation. In: Turner A.K., Schuster R.L. (Eds.), *Landslides: Investigation and Mitigation*. Special Report 247. Transportation Research Board, Washington D.C., 1996, 129-177
- [5] Guzzetti F., Carrara A., Cardinali M., Reichenbach P., Landslide hazard evaluation: A review of current techniques and their application in a multi-scale study, *Central Italy*. *Geomorphology*, 1999, 31(1-4), 181-216, DOI: 10.1016/S0169-555X(99)00078-1
- [6] Crozier M.J., Glade T., *Landslide Hazard and Risk: Issues, Concepts and Approach*. In: Glade T., Anderson M., Crozier M.J. (Eds.), *Landslide Hazard and Risk*, Wiley, Chichester 2012, 1-40, DOI: 10.1002/9780470012659.ch1
- [7] Guzzetti F., Reichenbach P., Ardizzone F., Cardinali M., Galli M., Estimating the quality of landslide susceptibility models. *Geomorphology*, 2006, 81(1-2), 166-184, DOI: 10.1016/j.geomorph.2006.04.007
- [8] Milevski I., Dragicevic S., Georgievska A., GIS and RS-based modelling of potential natural hazard areas in Pehchevo municipality, Republic of Macedonia. *Zbornik radova Geografskog instituta "Jovan Cvijić" SANU*, 2013, 63(3), 95-108
- [9] Milevski I., Ivanova E., GIS- and RS-based modelling of potential natural hazard areas in mountains. Case study: Vlahina mountain. In: Koulov B., Zhelezov G. (Eds.), *Sustainable Mountain Regions: Challenges and Perspectives in Southeastern Europe*. Springer, Cham, 2016, 191-204, DOI: 10.1007/978-3-319-27905-3_14
- [10] Milevski I., Dragicevic S., Radevski I., GIS and Remote Sensing based natural hazard modelling of Kriva River catchment, Republic of Macedonia. *Zeitschrift fur Geomorphologie*, 2017, 61(Suppl. 2), 213-228, DOI: 10.1127/zfg_suppl/2016/0364
- [11] Peševski I., Jovanovski M., Papić J., Abolmasov B., Model for GIS landslide database establishment and operation in Republic of Macedonia. *Geologica Macedonica*, 2015, 29(1), 75-86
- [12] Peshevski I., Jovanovski M., Abolmasov B., Papić J.B., Đurić U., Marjanović M., Haque U., Nedelkovska N., Preliminary regional landslide susceptibility assessment using limited data. *Geologia Croatica*, 2019, 72(1), 81-92, DOI: 10.4154/gc.2019.03
- [13] Reichenbach P., Rossi M., Malamud B., Mihir M., Guzzetti F., A review of statistically-based landslide susceptibility models. *Earth-Science Reviews*, 2018, 180, 60-91, DOI: 10.1016/j.earscirev.2018.03.001
- [14] Pourghasemi H.R., Teimoori Yansari Z., Panagos P., Pradhan B., Analysis and evaluation of landslide susceptibility: a review on articles published during 2005–2016 (periods of 2005–2012 and 2013–2016). *Arab Journal of Geosciences*, 2018, 11, DOI:

- 10.1007/s12517-018-3531-5
- [15] Feizizadeh, B., Blaschke, T., An uncertainty and sensitivity analysis approach for GIS-based multicriteria landslide susceptibility mapping. *International Journal of Geographical Information Science*, 2014, 28(3), 610-638, DOI: 10.1080/13658816.2013.869821
- [16] Chalkias C., Polykretis C., Ferentinou M., Karymbalis E., Integrating expert knowledge with statistical analysis for landslide susceptibility assessment at regional scale. *Geosciences*, 2016, 6(14), 1-15, DOI: 10.3390/geosciences6010014
- [17] Tošič R., Dragičević S., Zorn M., Lovrić N., Landslide susceptibility zonation: A case study of the municipality of Banja Luka (Bosnia and Herzegovina). *Acta geographica Slovenica*, 2014, 54(1), 189-202, DOI: 10.3986/AGS54307
- [18] Chalkias C., Ferentinou M., Polykretis C., GIS-Based landslide susceptibility mapping on the Peloponnese Peninsula, Greece. *Geosciences*, 2014, 4(3), 176-190, DOI: 10.3390/geosciences4030176
- [19] Donati L., Turrini M.C., An objective method to rank the importance of the factors predisposing to landslides with the GIS methodology: Application to an area of the Apennines (Valnerina; Perugia, Italy). *Engineering Geology*, 2002, 63(3-4), 277-289, DOI: 10.1016/S0013-7952(01)00087-4
- [20] Lima P., Steger S., Glade T., Tilch N., Schwarz L., Kociu A., Landslide Susceptibility Mapping at National Scale: A First Attempt for Austria. In: Mikos M., Tiwari B., Yin, Y., Sassa, K. (Eds.), *Advancing Culture of Living with Landslides: Volume 2, Advances in Landslide Science*. Springer, Cham, 2017, 943-951, DOI: 10.1007/978-3-319-53498-5_107
- [21] Komac B., Zorn M., Statistical landslide susceptibility modeling on a national scale: The example of Slovenia. *Revue Roumaine de Géographie*, 2009, 53(2), 179-195
- [22] Komac M., Ribičič M., Landslide susceptibility map of Slovenia at scale 1:250,000. *Geologija*, 2006, 49(2), 295-309, DOI: 10.5474/geologija.2006.022
- [23] Milevski I., Radevski I., Dimitrovska O., Svemir G., Digital model of the mean annual temperature and precipitation in Macedonia. *Geographical Reviews*, 2015, 48, 27-34
- [24] Mărgărint M.C., Grozavu A., Patriche C.V., Assessing the spatial variability of coefficients of landslide predictors in different regions of Romania using logistic regression. *Natural Hazards and Earth System Sciences*, 2013, 13(12), 3339-3355, DOI: 10.5194/nhess-13-3339-2013
- [25] Jaafari A., Najafi A., Pourghasemi H.R., Rezaeian J., Sattarian A., GIS-based frequency ratio and index of entropy models for landslide susceptibility assessment in the Caspian forest, northern Iran. *International Journal of Environmental Science and Technology*, 2014, 11(4), 909-926, DOI: 10.1007/s13762-013-0464-0
- [26] He Y., Beighley R.E., GIS-based regional landslide susceptibility mapping: A case study in southern California. *Earth Surface Processes and Landforms*, 2008, 33(3), 380-393, DOI: 10.1002/esp.1562
- [27] Saaty T.L., *The Analytical Hierarchy Process*. McGraw Hill, New York, 1980
- [28] Komac M., Landslide susceptibility model using the Analytical Hierarchy Process method and multivariate statistics in perialpine Slovenia. *Geomorphology*, 2006, 74(1-4), 17-28, DOI: 10.1016/j.geomorph.2005.07.005
- [29] Saaty T.L., Vargas L.G., *Models, Methods, Concepts, and Applications of the Analytic Hierarchy Process*. Springer, New York, 2012, DOI: 10.1007/978-1-4614-3597-6
- [30] Tanaka Y., Differences of landslide occurrences behavior due to slope aspects in the Amehata River Basin. *American Geophysical Union Fall Meeting*, 2005
- [31] Wu C., Ding M., Wang Q., Qiao J., Correlation between landslides and gradient in the Three Gorges Reservoir area based on GIS and information value model. In: Liu Y., Chen A. (Eds.), *18th International Conference on Geoinformatics*. Beijing, 2010, 1-5, DOI: 10.1109/GEOINFORMATICS.2010.5567596
- [32] Goepel K.D., Implementing the Analytic Hierarchy Process as a Standard Method for Multi-Criteria Decision Making in Corporate Enterprises – A New AHP Excel Template with Multiple Inputs. In: *Proceedings of the International Symposium on the Analytic Hierarchy Process*. Kuala Lumpur, 2013, DOI: 10.13033/isahp.y2013.047
- [33] Jovanovski M., Milevski I., Papić J.B., Peševski I., Markoski B., Landslides in Macedonia triggered by extreme events in 2010. In: Loczy D. (Ed.), *Geomorphological Impacts of Extreme Weather: Case Studies From Central and Eastern Europe*. Springer, Dordrecht, 2013, 265-279, DOI: 10.1007/978-94-007-6301-2_17
- [34] Fawcett T., An introduction to ROC analysis. *Pattern Recognition Letters*, 2006, 27(8), 861-874, DOI: 10.1016/j.patrec.2005.10.010
- [35] Lovrić N., Tošič R., Validation of landslide susceptibility maps (case study: urban area of the municipality of Banja Luka - B&H). *Glasnik Srpskog geografskog društva*, 2017, 97(1), 19-34, DOI: 10.2298/GSGD1701019L
- [36] Wang L.J., Guo M., Sawada K., Lin J., Zhang J., A comparative study of landslide susceptibility maps using logistic regression, frequency ratio, decision tree, weights of evidence and artificial neural network. *Geosciences Journal*, 2016, 20(1), 117-136. DOI: 10.1007/s12303-015-0026-1
- [37] Saro L., Woo J., Kwan-Young O., Mounj-Jin L., The spatial prediction of landslide susceptibility applying artificial neural network and logistic regression models: A case study of Inje, Korea. *Open Geosciences*, 2016, 8(1), 117-132. DOI: 10.1515/geo-2016-0010
- [38] De Graff J.V., Romesburg H.C., Ahmad R., McCalpin J.P., Producing landslide-susceptibility maps for regional planning in data-scarce regions. *Natural Hazards*, 2012, 64(1), 729-749, DOI: 10.1007/s11069-012-0267-5
- [39] Zorn M., Komac B., Kumelj Š., Mass movement susceptibility maps in Slovenia: The current state. *Geografski vestnik*, 2012, 84(1), 99-112
- [40] Zorn M., Komac B., Natural disasters and social irresponsibility. *Geografski vestnik*, 2015, 87(2), 75-93, DOI: 10.3986/GV87205