Landslide Hazard Zonation Using Quantitative Methods in GIS

M. H. Vahidnia[1], A. A. Alesheikh[2,*], A. Alimohammadi[3], F. Hosseinali[4]

Received: November 2008 Accepted: June 2009

Abstract: *Landslides are major natural hazards which not only result in the loss of human life but also cause economic burden on the society. Therefore, it is essential to develop suitable models to evaluate the susceptibility of slope failures and their zonations. This paper scientifically assesses various methods of landslide susceptibility zonation in GIS environment. A comparative study of Weights of Evidence (WOE), Analytical Hierarchy Process (AHP), Artificial Neural Network (ANN), and Generalized Linear Regression (GLR) procedures for landslide susceptibility zonation is presented. Controlling factors such as lithology, landuse, slope angle, slope aspect, curvature, distance to fault, and distance to drainage were considered as explanatory variables. Data of 151 sample points of observed landslides in Mazandaran Province, Iran, were used to train and test the approaches. Small scale maps (1:1,000,000) were used in this study. The estimated accuracy ranges from 80 to 88 percent. It is then inferred that the application of WOE in rating maps' categories and ANN to weight effective factors result in the maximum accuracy.*

Keywords: *Landslide Susceptibility Map, Artificial Neural Network, Weight-of-Evidence, Analytical Hierarchy Process, General Linear Regression*

1. Introduction

Landslides are significant natural geologic hazard around the world. Expansion of urban and man-made structures into potentially hazardous areas leads to extensive damage to infrastructure and occasionally results in loss of life every year.

Since the early 1970s, many scientists have attempted to assess landslide hazards and produced susceptibility maps portraying their spatial distribution by applying many different GIS based methods. The results of published papers show that landslide susceptibility and hazard maps have become very effective tools for planners and decision makers [1].

Landslide hazard (LH) defines the physical attributes of a potentially damaging landslide in terms of mechanism, volume and frequency [2] and therefore landslide hazard assessment (LHA) estimates the probability of a landslide occurrence within a certain period of time in a given area [3] [4]. Most attempts have been made to establish the intensity and dimension of landslides compared with temporal frequency of slope failures [5]. Such kind of LHA is expressed in the term of Landslide Susceptibility Assessment (LSA). Hence, a hazard map that aims at predicting where slope failures are most likely to occur, is more accurately defined as a landslide susceptibility map. The term susceptibility defines the likelihood of occurrence of a landslide if governing factors like rainfalls, earthquakes, etc., are not considered [6]. According to the Soriso Valvo (2002) [5] landslide susceptibility assessment is a process directed to establish the likelihood that future landslides will occur in a given area on the basis of suitable physical terrain factors (slope, land use, litholgy, etc.) In LSA, spatial correlation between the important terrain factors and past landslide distribution is often considered [7]. An up-to-date landslide inventory map that includes existing information on the past mass movement is then required to determine these correlations [3]. The reliability and accuracy of the future spatial probability of slope failure that is portrayed in Landslide Susceptibility Map (LSM) [8] must also be gathered. The accuracy mostly

* *Corresponding Author: Email:alesheikh@kntu.ac.ir, Tel: (+98 21)8878 6212, Fax: (+98 21) 8878 6213,*

1 *Graduate Student, Department of Geospatial Information System (GIS), vahidnia84@gmail.com*

2 *Associated Professor, Department of Geospatial Information System (GIS),*

3 *Assistant Professor, Department of Geospatial Information System (GIS), alimoh_abb@kntu.ac.ir*

4 *PhD Student, Department of Geospatial Information System (GIS), frdhal@gmail.com*

Faculty of Geodesy and Geomatics Eng. K.N.Toosi University of Technology, No 1346, Valiasr Street, Mirdamad Cross, Tehran, IRAN 19967-15433

depends on the amount and quality of available data, the working scale and the selection of the appropriate methodology of analysis and modeling [9] [10].

Different models and methods have been proposed to produce LSM; however, the general agreement on these methods is yet to come. Despite of these disagreements, all the methods can be classified into direct and indirect techniques or, into qualitative and quantitative approaches [11].

In a direct mapping approach the degree of hazard is mapped directly in the field, or is determined on the basis of a detailed geomorphological map. Despite of some advantages of direct method, it is, however, a time-consuming method that depends on the expertise of geomorphologists [12].

The indirect method shares the following steps [13]: mapping of the landslides in the target region, mapping of a set of geological – geomorphological factors which are directly or indirectly correlated with slope instability, determinig the relative correlation with slope instability, and classifying the land surface into the regions of different landslide susceptibility degree (hazard zoning).

Qualitative methods are subjective and demonstrate the hazard zoning in descriptive (qualitative) terms. Quantitative methods produce numerical estimates (probabilities) of the occurrence of landslide phenomena in any hazard zone [14]. In general, a qualitative approach is based on the subjective judgment of an expert or a group of experts but the quantitative approach is based on mathematically objective structures [2].

Landslide susceptibility assessment was challenged by several researchers in recent years. The use of Analytic Hierarchy Process (AHP) by Komac [15]; Yalcin [9] and Analytic Network Process (ANP) [16] by Neaupane and Piantanakulchai [2], weighted linear combination (WLC) by Ayalew et al. [17] and Fuzzy Logic theory by Ercanoglu and Gokceoglu [18] are all different examples of qualitative or semi quantitative approaches in LSA. Also widely use of bivariate statistical models by Lin and Tung [19]; Su¨zen and Doyuran [13]; Thiery et al. [20] and various multivariate statistical techniques like discriminant analysis by Carrara et al. [21] or linear and logistic regression by Dai and Lee [22]; Ayalew and Yamagishi [10], Artificial Neural Network (ANN) by Yesilnacar and Topal [23]; Lee et al. [24]; Ermini et al. [25] are all different examples of quantitative techniques. These statistical and functional methods facilitate understanding of the relationships between landslides and preparatory factors, and guarantee lower subjectivity levels with respect to qualitative approaches and, therefore, have been attempted more than the others by academic and research institutions [24].

2. Experimental procedures

The study area is located in the south of Mazandaran province northern Iran where most landslides occurring in the country have been reported from this chiefly mountainous and forestry region. The area under investigation is located from 52°31′ to 53°27′ east longitude and from 35°52′ to 36°30′ north latitude with expansion of about 3440 km² and elevation ranges from 45 to 3273 m (Figure.1). The environmental setting, excessive rain, relatively extensive drainage network, low resistance of the soil and rocks against pressure and variation of slope are recognized as the most effective parameters of landslide and slope failure in this region. The main units of lithology in this region are consist of limestone, dolomite, shale,

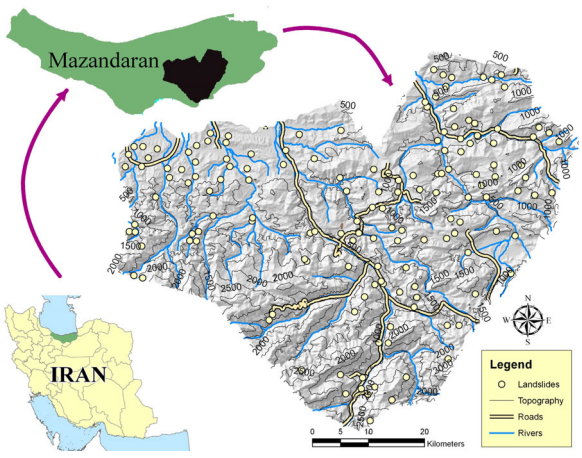

Fig. 1. Study area and landslide inventory map

siltstone, sandstone, marl, tuff, conglomerate, anhydrite, salt and their different combination or derivate. Also the main units of landuse of study area are forest with variant range of coverage from low to dense, garden, farming land, and their different combination or derivate, and a little urban and water regions. The major landslides in the region are transitional or rotational and seldom rockfall that will be modeled together in this study.

In this study, the susceptibility mapping started with the preparation of an inventory map of 151 landslides of the study area at the scale of 1:1,000,000. Seven causal factors are then chosen, namely; slope, aspect, curvature, land cover, lithology, distance to river and distance to fault. The important factors that are considered in the selection process where: having a certain degree of correlation with landside, being fairly represented all over the study area, having spatial variation, measurability and non redundancy [9]. All preliminary thematic maps (i.e. landslides inventory, and land cover from Jihad-e-Keshavarzi Ministry Digital Elevation Model (DEM) with accuracy of 1 meter, river, fault, and lithology of study area were obtained from National Geosciences Database of Iran (NGDIR) in small scale (1:1,000,000).

LSM is generated through various approaches; namely, Weight of Evidence (WOE), Artificial Neural Network (ANN), Analytical Hierarchy Process (AHP) and General Linear Regression (GLR). Next subsections describe these concepts and the steps to implement each method.

2.1. Weight of Evidence (WOE)

This method is based on Bayesian probability theory and was developed for the identification and exploration of mineral deposits [26]. In comparison with other analyses, such as multiple logistic regression analysis, discriminant analysis, factor analysis, and cluster analysis, the results of weights-of-evidence are easy to interpret. Moreover, spatial patterns with complex geometries can be modeled with the same computational effort as patterns with simple geometries. The effect of each spatial variable can also be calculated independently of a combined solution [26]. Prior and posterior probabilities are the major concepts which are used in this approach to determine the relative importance of data [27]. The probability P is usually estimated empirically with knowledge about the occurrence of an event D in the past under equal conditions, and is known as prior probability $P\{D\}$ where $P\{B\}$ denotes the evidence probability. When the evidences are integrated into the calculation of the probability, it is identified as conditional or posterior probability $P\{D|B\}$. This posterior probability $P\{D|B\}$ expresses the probability that an event D will occur under the presence of an evidence B. Both probabilities (prior and posterior) are integrated into the Bayes theorem as follows:

$$P\{D \mid B\} = \frac{P\{D\}.P\{B \mid D\}}{P\{B\}} \quad (1)$$

WOE has the ability of measuring the spatial relation between the landslides and the evidences named by W^+ and W^- and consequently their contrast C as:

$$C = W_j^+ - W_j^- = \ln\left(\frac{P\{D \mid B_i\}}{P\{D \mid \overline{B}_i\}} \Big/ \frac{P\{\overline{D} \mid B_i\}}{P\{\overline{D} \mid \overline{B}_i\}}\right) \quad (2)$$

where W^+ and W^- describe the probably that a landslide will occur in the case of presence or absence of evidence B respectively [26]. If the presence of evidence B in landslide occurrence is more effective than what would be expected by chance, W^+ is positive and W^- is negative. Conversely, if the presence of evidence B in landslide occurrence is less effective than what would be expected by chance, W^+ is negative and W^- is positive. The contrast C represents the extent of these spatial associations. A larger C value indicates a strong spatial association between the occurrences and the evidence map [27].

In this study, each category of causal factors was used as evidence and the contrast values (C) was calculated from the WOE method. In order to prevent undesirable effects of negative weights and probability of zero in results (i.e. $Ln\ 0 \rightarrow -\infty$ while $e^{-\infty} \rightarrow 0$), the quantity e^C was used to rate the classes in the first (WOE-ANN) and second (WOE-ANN-AHP) approaches.

2.2. Analytical Hierarchy Process (*AHP*)

AHP is a well-known multi-attribute weighting method for decision making. Pairwise comparisons are used in this decision-making process to form a reciprocal matrix by transforming qualitative data to crisp ratios. The reciprocal matrix is then solved by a weight finding method for determining the criteria importance and alternative performance [28]. Once the pairwise comparison matrix is obtained; based on the problem elements, the aim is to summarize preferences so that each element can be assigned a relative importance [29]. The Eigenvalue method is one way to access ultimate weights of criteria. In order to measure the degree of consistency of decision making, the consistency index (*CI*) is estimated through:

$$CI = \frac{\lambda_{\max} - p}{p - 1} \tag{3}$$

where λ_{max} is the biggest eigen value of comparison matrix and p is the number of criteria or dimension of matrix. The consistency ratio (*CR*) can be computed as:

$$CR = \frac{CI}{RI} \tag{4}$$

where *RI* is the random index (the consistency index of a randomly generated pairwise comparison matrix). It can be shown that *RI* depends on the number of elements being compared.

The consistency ratio (*CR*) is designed in such a way that if *CR*<0.10 then the ratio indicates a reasonable level of consistency in the pairwise comparison; if, however, *CR*>=0.10, then the values of the ratio are inconsistent [30].

In this study the AHP comparison matrix was used to achieve causative factor maps' weights and verify the consistency of relative weights resulting from a data-driven method (ANN) and in a pairwise comparison of criteria. The AHP method usually can estimate better results than other simple knowledge-based weighting methods especially in landslide problems [9].

2.3. Artificial Neural Network (ANN)

Artificial Neural Networks are computational networks which attempt to simulate, the networks of nerve cell (neurons) of the biological central nervous system [31]. It has been found that ANNs, specifically Multilayer Perceptron (MLP) model, have several advantages for Landslide Susceptibility mapping, such as the ability to handle imprecise and fuzzy information, fault and failure tolerance, high parallelism, non-linearity, robustness, capability to generalize and tolerance to noise data and thus have the capability to analyze complex data patterns [32]. Ability of learning is one of the most important characteristics of ANN [26]. It means that the random initialization of the weights of network at first would be adjusted according to the different patterns of data. Compared with the statistical methods, neural networks need less training data for accurate analysis [33]. A MLP consists of three layers namely; input, output and hidden layers [34] as represented in Figure.2.

This network consists of a number of interconnected nodes from all layers. Each node is a simple processing element that responds to the weighted inputs which is received from other nodes. The receiving node sums the weighted signals from all nodes to which it is connected in the preceding layer [35]. The number of neurons (nodes) in the input layer is equal to the number of data sources and the number of neurons in the output layer is constrained by the application and is represented by the number of outputs. The number of hidden layers and the number of neurons in each layer depends on the architecture of network and usually are determined by trial and error [33]. The network weights are then modified in the training process by a number of learning algorithms based on back propagation learning [26]. In order to solve Non-linear classification problems and to adjust weights of hidden layer in back propagation learning any

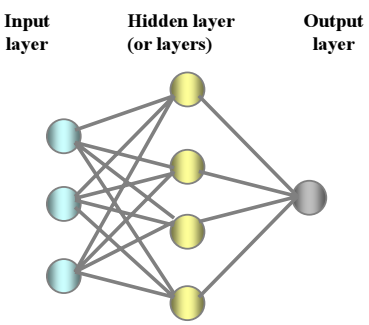

Fig. 2. The structure of Multilayer Perceptron with one hidden layer

differentiable nonlinear function can be used as an activation function, but a sigmoid function such as equation 1 is generally used [26]:

$$f(net_{pj}) = \frac{1}{1+e^{-net_{pj}}} \quad (5)$$

where net_{pj} is the weighted sum of pattern p in node j. The error of pattern p in node j (i.e. δ) for output layer and hidden layer are obtained respectively from the following equations:

$$\delta_{pj} = o_{pj}(1-o_{pj})(t_{pj}-o_{pj})$$
$$\delta_{pj} = o_{pj}(1-o_{pj})\sum_{k}\delta_{pk}w_{jk} \quad (6)$$

Where o_{pj} is the real output of pattern p in node j, t_{pj} is the target output of pattern p in node j, w_{jk} is the weight between node j and k in next layer and $\sum$ operator is related to k units in layer after node j.

2.4. General Linear Regression (GLR)

Linear regression is a form of multivariate statistical analysis in which observational data are modeled by a function. The function has usually assumed one response variable Y and m predictor variables $X1, X2, \ldots, Xm$ and n observations [29]. In this case study Y is the presence or absence (or the area percentage) of landslides in each sampling unit, the X's are input predictor variables (the area of categories in causative factors) measured or observed for each mapping unit, the β's are coefficients estimated from the data and ε represents the model error. In matrix terms this becomes:

$$Y = X\beta + \varepsilon \quad (7)$$

According to the principle of least squares, the fitted values of β that minimize ε are:

$$\beta = (X'X)^{-1}X'Y \quad (8)$$

In this case the residuals are defined as the difference between observed and fitted values and so there are some criteria to evaluate the correctness of fitted model that will be mentioned as: Sum of squared errors related to regression (SS_{reg}) and Sum of squared errors related to residuals (SS_{res}). Then R^2 factor would be acquired as:

$$R^2 = \frac{SS_{reg}}{SS_{res}+SS_{reg}} \quad (9)$$

R^2 is used to verify the efficiency of model (if R^2 approximately equals 1) and also interpreted as the amount of variability in the observations that can be explained by the predictors. This model will act well for landslide assessment if the problem have a near to linear structure.

Also, the sample variance and Mean Square of the observations (MS_{reg}) and residuals (MS_{res}), by using the regression equation, according to the number of equations (n) and unknown parameters (m) are respectively given by:

$$MS_{reg} = \frac{SS_{reg}}{m} \quad (10)$$

$$MS_{res} = \frac{SS_{res}}{n-m-1} \quad (11)$$

The t-distribution and Fisher distribution tests may be used for testing the outcomes [29].

3. Methodologies and Results

In order to assess the landslide susceptibility, four approaches were practiced that are mainly different in the process of weighting or rating. The results of each approach in weighting phase are values representing the relative importance of factors and their categories respectively. These approaches were applied and then compared.

3.1. First Approach (WOE-ANN)

The first approach uses the Weight Of Evidence (WOE), for rating categories of each causative factor map of instability, and an Artificial Neural Network (ANN), for weighting factor maps, as represented in Figure.3.

By overlaying landslide inventory map with each causative factor, the statistical relationship can be measured between categories and past landslides based on landslide density in each category, and assessed Contrast values (C) as the rating values. The results of rating step are shown in Table 1.

In order to calculate relative weights between the factor maps, a multilayer perceptron was

Table 1 The results of rating by WOE in the first approach

Effective factors	Categories	C	e^C	Effective factors	categories	C	e^C
Lithology	Jd	0.5930	1.8094	Land Cover	mix(agri_garden-lowforst)	0.2666	1.3056
	Jk	0.7496	2.1162		mix(dryfarming_lowforest)	-∞	0
	Czl	-∞	0		mix(agri_follow)	-0.0714	0.9310
	Jl	-1.5707	0.2078	Dist. to river	0-500 m	0.3190	1.3758
	Db.sh	-∞	0		500-1000 m	0.0381	1.0389
	Pr	-∞	0		1000-1500 m	-0.0284	0.9719
	TRe	-0.6149	0.5406		1500-2000 m	0.2781	1.3207
	Cb	-∞	0		>2000 m	-0.5465	0.5789
	Ktzl	1.1242	3.0778	Dist. to fault	0-1000 m	0.1146	1.1214
	Ek	-0.1882	0.8284		1000-2000 m	-0.0440	0.9569
	K1bvt	-∞	0		2000-3000 m	-0.3051	0.7370
	Plc	-0.1851	0.8309		3000-4000 m	0.1829	1.2007
	Mm,s,l	-∞	0		>4000 m	0.0262	1.0265
	Qft2	-∞	0	Slope	0° - 5°	-0.7038	0.4946
	Mlgs	0.3380	1.4021		5° -15°	0.1455	1.1567
Land Cover	mix(lowforest_goodrange)	-0.6741	0.5095		15° -25°	-0.0010	0.9989
	Modforest	-0.5187	0.5952		25° -35°	-0.0706	0.9317
	Garden	-∞	0		>35°	0.6136	1.8472
	mix(modforest_goodrange)	-1.1177	0.3270	Aspect	North	-0.5297	0.5887
	Goodrange	-0.8417	0.4309		North East	0.2640	1.3021
	mix(goodrange_garden)	-0.0471	0.9539		East	0.0121	1.0121
	Lowforest	-0.4621	0.6299		South East	0.0344	1.0350
	mix(garden_lowforest)	0.2602	1.2972		South	0.3497	1.4187
	Urban	1.7238	5.6060		South West	0.1315	1.1405
	mix(lowforest_agri)	-∞	0		West	-0.5755	0.5623
	mix(modrange_modforest_garden)	-∞	0		North West	0.1377	1.1476
	Denseforest	0.3773	1.4584	Curvature	Very concave	-0.0103	0.9896
	mix(dryfarming_follow_modforest)	-0.1511	0.8597		Concave	0.0591	1.0608
	mix(agri_goodrange)	-∞	0		Gently convex	-0.0886	0.9151
	Water	-∞	0		Convex	0.1103	1.1166
	mix(dryfarming_follow)	-∞	0		Very convex	-0.1454	0.8646
	Agri	-∞	0				

established. The weights were acquired then by training of the neural network with a training set that consists of landslide points (sliding conditions) and non-landslide points (stability conditions) and their associated values in each factor map as input values.

After using a test set and calculating RMSE for different network structures, the network structure associated with the minimum RMSE was selected as the best network structure. This

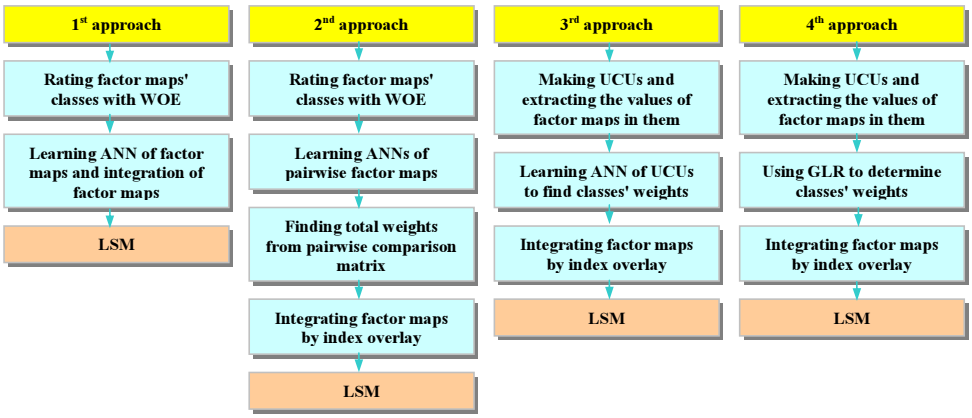

Fig. 3. The general process of each approach in this study

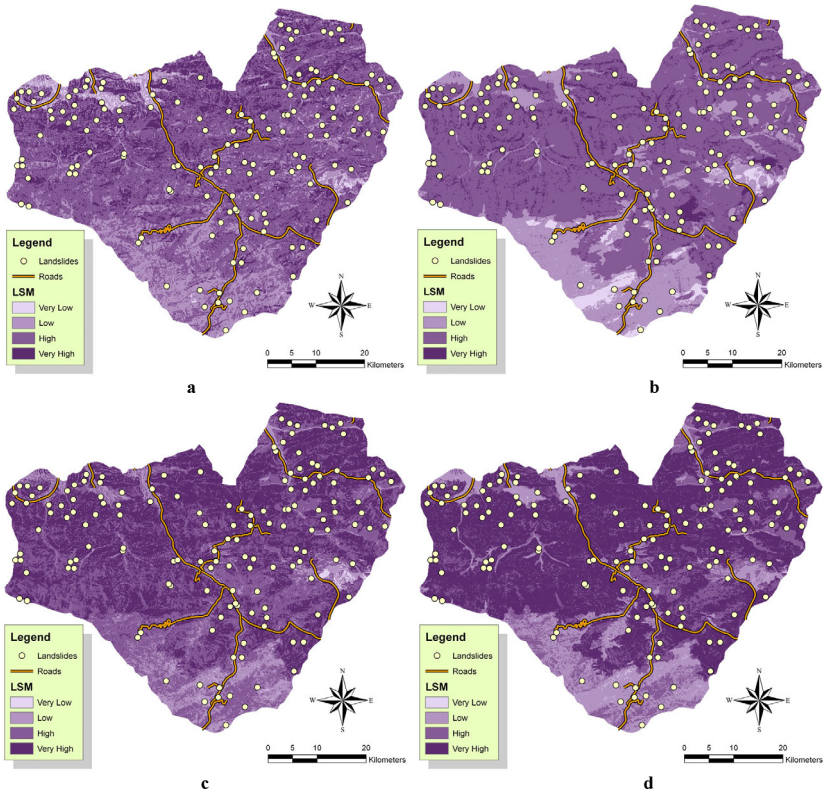

Fig. 4. Landslide Susceptibility Map (LSM) from four methods: a-1st approach (using Weight-of-Evidence and Artificial Neural Network), b-2nd approach (using Weight-of-Evidence, Artificial Neural Network and Analytical Hierarchy Process), c-3rd approach (using Unique Condition Units in Artificial Neural Network), d-4th approach (using Unique Condition Units in Generalized Linear Regression).

structure was 7×4×2×3×1 (i.e. 7 neurons in input layer, 4 neurons in the first hidden layer, 2 neurons in the second hidden layer, 3 neurons in the third hidden layer and 1 neuron in output layer). A MATLAB-based application program was developed and their result is shown in Table 2. Then all factor maps were entered into the optimum MLP structure for integrating and were reclassified into 4 domains of Very High Risk, High Risk, Low Risk and very Low Risk in order to represent different degree of hazard at the study area (Figure.4-a).

3.2. Second Approach (WOE-ANN-AHP)

The rating phase in this approach is the same as the first approach. But, in weighting phase, each pair of factor maps is applied in a MLP for

Table 2 Optimized network structure in the first approach

Number of neuron in each layer					RMSE of training set
Input layer	1st hidden layer	2nd hidden layer	3rd hidden layer	Output layer	
7	4	2	3	1	0.4079

Table 3 Optimized network structure of each pairwise comparison of factor maps and their relative weights in the second approach

Comparison factors		Number of neuron in each layer					weights
1st factor	2nd factor	Input Layer	1st Hidden layer	2nd Hidden layer	3rd Hidden layer	Output layer	
Slope	Aspect	2	5	7	1	1	0.624
Slope	Curvature	2	5	7	8	1	0.994
Slope	Landcover	2	8	7	8	1	0.912
Slope	Lithology	2	3	-	-	1	1.004
Slope	Dist. to river	2	4	8	1	1	0.998
Slope	Dist. to fault	2	8	3	7	1	0.985
Aspect	Curvature	2	8	3	2	1	0.977
Aspect	Landcover	2	6	8	8	1	0.990
Aspect	Lithology	2	6	-	-	1	0.153
Aspect	Dist. to river	2	7	-	-	1	0.992
Aspect	Dist. to fault	2	7	7	6	1	0.971
Curvature	Landcover	2	4	-	-	1	0.950
Curvature	Lithology	2	5	1	5	1	1.027
Curvature	Dist. to river	2	2	4	1	1	1.016
Curvature	Dist. to fault	2	4	6	7	1	4.496
Landcover	Lithology	2	1	1	-	1	1.359
Landcover	Dist. to river	2	8	8	7	1	0.990
Landcover	Dist. to fault	2	6	-	-	1	0.997
Lithology	Dist. to river	2	3	6	3	1	0.966
Lithology	Dist. to fault	2	6	4	2	1	1.036
Dist. to river	Dist. to fault	2	3	3	7	1	0.995

training of the network with training data set and controlling with testing data set. Ultimately, the weight of each pair in landslide occurrence was extracted (Figure.3). The number of factor maps would be 21 since there exists 7 networks needed for all cases of pairwise factor maps (with elimination of extra or repeated cases). The network output, like the previous approach, shows the degree of landslide presence in the study area. Table 3 represents the optimized network of each pairwise factor maps and their relative weights.

These weights are then formed a pairwise comparison matrix of AHP. Final weights of causative factor maps were calculated from eigenvalue method and the Consistency Ratio are represented in Table 4. The Consistency Ratio is acceptable and therefore the results are confirmable.

The factor maps were integrated in this approach by using Index Overlay. The reclassified map of integration step would be a landslide susceptibility map as seen in Figure 4-b.

3.3. Third Approach (UCU-ANN)

This approach is used homogeneous domains called Unique-condition units [25] as the cell unit. The UCU are singled out by sequentially overlying all the layers, whose number, size and

Table 4 Final weights of factor maps resulted from an AHP pairwise comparison matrix with an acceptable Consistency Ratio in the second approach

Factors	Normalized weights	Factors	Normalized weights
Slope	1.212	Lithology	2.042
Aspect	1.147	Dist. to river	1.295
Curvature	1.789	Dist. to fault	1.107
Landcover	1.408		
λmax	7.7715	**CR**	0.0974

Table 5 Optimized network structure in the third approach

Properties	Value	Properties	Value
Neurons in input layer	63	Number of input patterns	32870
Neurons in hidden layer	12	RMSE of training set	0.946
Neurons in output layer	1	RMSE of testing set	1.013

nature depend on the criteria used in classifying the input factors [13]. In this approach rating and weighting was done simultaneously for all categories of all causative factor maps in one neural network (Figure 3); therefore, there are 63 neurons (number of categories) in input layer of multilayer perceptron. It seems that when statistical and functional methods are used, all the categories in all thematic layers would be comparable.

In order to reach this objective a UCU layer was established over the study area. This layer divides the study area into about 33000 homogeneous domains. Next step was devoted to determine the presence or absence of each category in each unit which assigned values 1 or 0. These values form inputs of a MLP with output values of 1 or 0 associated to the case of landslide or non landslide in each unit. Table 5 shows the structure of optimized network that is 63×12×1. The final weights are extracted from multiplying matrices of 63×12 and 12×1 dimensions (Table 6). After the combining factor maps by Index Overlay and reclassification of the output, the LSM was obtained as represented in Figure 4-c.

3.4. Fourth Approach (UCU-GLR)

In this approach, like the third approach, Unique-condition units (UCUs) were used as the cell unit and reminder was related to determine presence or absence of each class and landslide occurrence in these units as done in the previous approach (Figure 3). The presence or absence of landslide is a dependent variable and categories of parameters are independent variables which form the matrices of GLR. The coefficients of best fitted linear model would be known as the categories' weights consequently. The results of regression analysis and general testing of model are represented in Table 7; the selected predictors and their associated coefficients are represented in Table 8.

As seen 59 predictors from all 63 primary parameters have been selected with the Confidence Interval more than 95 percent. This model was then applied to integrate all classes of the effective parameters (Figure 3) however the linear dependency in this case is not completely significant according to the R^2 testing result. Figure 4-d shows the LSM concluded from this approach after applying predictors' weights, overlying factor maps and then reclassifying the output map based on the degree of susceptibility.

4. Conclusion and Discussion

This paper tried to generate Landslide Susceptibility Map (LSM) using various approaches. These approaches include statistical methods that consist bivariate statistical models, like WOE, and multivariate statistical models, like GLR, Intelligent method with emphasis on Multilayer Perceptron (MLP) as a common ANNs and also Pairwise Comparison Matrix (PCM) of AHP as a decision making tool in a data driven manner. The paper classified all the methods in four categories; namely; WOE-ANN,

Table 6 Weights of categories in the third approach

factors	Classes	Weight	factors	Classes	factors
Lithology	Jd	0.1521	Land Cover	mix(agri_garden-lowforest)	0.2254
	Jk	1.3520		mix(dryfarming_follow_lowforest)	-0.2597
	Czl	0.0577		mix(agri_follow)	0.0729
	Jl	1.0103	Distance to river	0-500 m	-0.4185
	Db.sh	0.7769		500-1000 m	-0.2537
	Pr	1.5422		1000-1500 m	0.7209
	TRe	0.7218		1500-2000 m	0.2016
	Cb	0.0290		>2000 m	-0.0884
	Ktzl	-0.8631	Distance to fault	0-1000 m	1.1854
	Ek	0.4531		1000-2000 m	-0.6967
	K1bvt	0.9246		2000-3000 m	0.7809
	Plc	-0.0045		3000-4000 m	-0.4821
	Mm,s,l	-0.7007		>4000 m	-0.3157
	Qft2	-0.1792	Slope	0° - 5°	0.4426
	Mlgs	-0.6582		5° -15°	0.1523
Land Cover	mix(lowforest_goodrange)	-0.6343		15° -25°	-0.6453
	modforest	-0.4369		25° -35°	0.0045
	garden	-0.4091		>35°	-0.5316
	mix(modforest_goodrange)	-0.2237	Aspect	North	-0.1696
	goodrange	-0.8735		North East	0.8314
	mix(goodrange_garden)	-0.3737		East	-0.6595
	lowforest	-0.4025		South East	0.5389
	mix(garden_lowforest)	-0.2480		South	0.7729
	urban	-0.6226		South West	0.1023
	mix(lowforest_agri)	0.0704		West	0.4682
	mix(modrange_modforest_garden)	0.4326		North West	0.4018
	denseforest	0.2084	Curvature	Very concave	-3.0984
	mix(dryfarming_follow_modforest)	0.0449		Concave	0.2168
	mix(agri_goodrange)	-0.5959		Gently convex	-0.5196
	water	1.7066		Convex	-0.9365
	mix(dryfarming_follow)	-0.6031		Very convex	-0.6757
	agri	-0.3032			

Table 7 The properties of Generalized Linear Regression model in the fourth approach

Model Summary					
		Number of observation (n)	**R**	**R^2**	**Std. Error**
		33237	.129	0.017	0.066
Analysis Of Variance (ANOVA)					
		Sum of Squares (SS)	**df**	**Mean Square (MS)**	**P ($F<F_{59,33177,0.05}$)**
	Regression	2.441	59	0.041	9.54
	Residual	143.908	33177	0.004	
	Total	146.35	33236		

Table 8 Selected coefficients of Generalized Linear Regression model and their related tests in the fourth approach

Predictors and constant value	B	Std. Err	t	P(|t| <$t_{n,0.025}$)	Predictors and constant value	B	Std. Err	t	P(|t| <$t_{n,0.025}$)
(Constant)	-0.007	0.003	-2.708	0.007	mix(agri_goodrange)	0.000	0.003	0.178	0.859
aspect E	0.005	0.001	4.704	0.000	water	0.002	0.006	0.265	0.791
aspect N	0.007	0.001	6.546	0.000	mix(dryfarming_follow)	0.000	0.008	-0.089	0.929
aspect NE	0.005	0.001	5.031	0.000	agri	-0.001	0.01	-0.141	0.888
aspect NW	0.007	0.001	6.402	0.000	mix(agri_garden-lowforest)	0.005	0.003	1.635	0.102
aspect S	0.006	0.001	5.437	0.000	mix(dryfarming_follow_lowforst)	0.003	0.012	0.224	0.823
aspect SE	0.006	0.001	6.006	0.000	mix(agri_follow)	0.004	0.003	1.545	0.122
aspect SW	0.006	0.001	5.314	0.000	Lithology Jd	-0.005	0.003	-1.857	0.063
aspect W	0.004	0.001	3.3	0.001	Lithology Jk	-0.008	0.003	-3.109	0.002
curve very concave	-0.003	0.001	-2.291	0.022	Lithology Czl	0.002	0.007	0.225	0.822
curve gently convex	0.002	0.001	2.134	0.033	Lithology Jl	-0.005	0.001	-3.901	0.000
curve convex	0.000	0.001	-0.363	0.717	Lithology Db.sh	0.000	0.006	0.065	0.949
curve very convex	-0.003	0.001	-2.745	0.006	Lithology Pr	0.000	0.006	-0.106	0.916
fault_0_1000	0.001	0.001	1.141	0.254	Lithology Tre	-0.004	0.002	-2.091	0.036
fault_2000_3000	0.000	0.001	-0.752	0.452	Lithology Cb	0.004	0.011	0.387	0.699
fault_3000_4000	-0.001	0.001	-0.858	0.391	Lithology Ktzl	-0.004	0.003	-1.397	0.163
fault_more_than_4000	0.001	0.001	0.941	0.347	Lithology Ek	-0.005	0.001	-3.628	0.000
mix(lowforest_goodrange)	0.000	0.003	0.066	0.948	Lithology K1bvt	-0.005	0.004	-1.37	0.171
modforest	0.002	0.002	0.904	0.366	Lithology Plc	-0.007	0.002	-4.493	0.000
garden	0.002	0.004	0.531	0.595	Lithology Mm,s,l	-0.01	0.003	-3.257	0.001
mix(modforest_goodrange)	0.002	0.003	0.705	0.481	Lithology Qft2	-0.006	0.003	-1.95	0.051
goodrange	0.004	0.003	1.44	0.150	river_0_500	0.003	0.001	2.139	0.032
mix(goodrange_garden)	0.002	0.002	0.828	0.408	river_500_1000	0.002	0.001	1.974	0.048
lowforest	0.003	0.002	1.214	0.225	river_1000_1500	0.001	0.001	0.639	0.523
mix(garden_lowforest)	0.004	0.003	1.392	0.164	river_1500_2000	0.001	0.001	1.137	0.256
urban	0.011	0.004	2.511	0.012	river_more_then_2000	0.001	0.001	0.626	0.531
mix(lowforest_agri)	0.000	0.004	-0.087	0.931	slope_0_5	-0.009	0.001	-6.815	0.000
mix(godrange_modforest_garden)	0.000	0.01	-0.057	0.955	slope_5_15	0.001	0.001	0.678	0.498
denseforest	0.011	0.002	5.905	0.000	slope_25_35	-0.002	0.001	-1.586	0.113
mix(dryfarming_follow_modforst)	0.006	0.002	2.743	0.006	slope_more_than_35	-0.003	0.002	-1.576	0.115

WOE-ANN-AHP, UCU-ANN, and UCU-GLR. Moreover, this research focused on the combination of techniques and the strength of ANN with respect to other regression models. Using the methods, it is shown that the meaningful intervals; which were computed by WOE method, facilitate the process of learning in ANN. This issue had been missed in previous studies. The paper also showed how to exercise a quantitative AHP rather than a common qualitative one as it is done in most studies. According to the type of data used, small scale maps in this study, the application of UCUs faced some troubles due to the exceeding volume of input patterns. Therefore, the cell units were more efficient than UCUs. This is another contribution of the paper. As discussed, all of these methods are quantitative without taking any expert's opinion into account and in all approaches the most important issue is devoted on investigation of correlation and dependency between causative factors or their classes and previous slope failures and then passing among rating and weighting phase. All the LSM of these

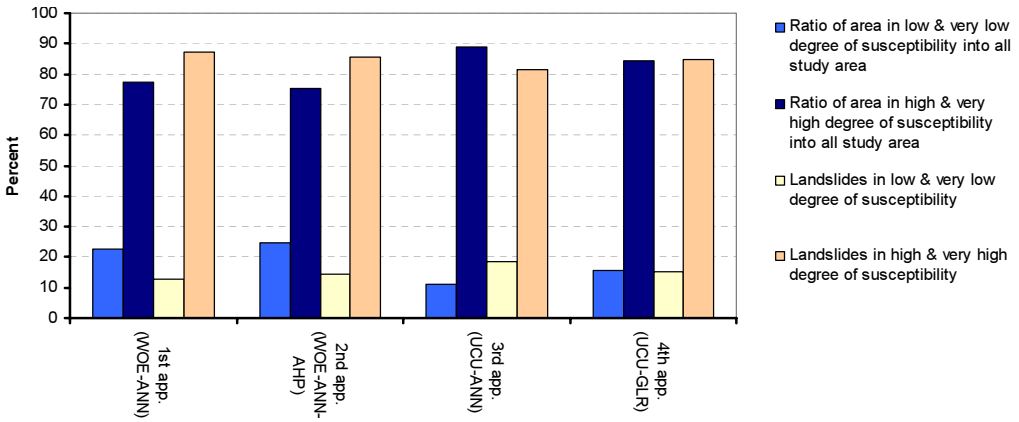

Fig. 5. Comparison of different approaches with considering landslide inventory map

4 approaches were consequently obtained from an equal interval classification that maps the results into four categories with different degree of susceptibility or risk such as: Very Low, Low, High and Very High (Figure 4). A quantitative validation of how ANNs performs in the landslide susceptibility mapping is made by comparing the areas predicted as High and Very High risk with the landslide inventory map. Comparison of these approaches based on this axiom (Figure 5) show that the 1st approach (WOE-ANN), by including 88 percent of landslide occurrence in High and Very High susceptibility classes, has been the best landslide estimator in this study, although all the accuracies in different approaches are in a reasonable range between 80 and 88 percent. The verification results showed a satisfactory agreement between the susceptibility map and landslide data. Numerous methods are proposed to compare and validate the results; some of them consider areas of hazard groups. In this paper, the overall accuracy has also been estimated by considering the area of each hazard zone and the percentage of its landslides frequency. Therefore, a higher accuracy can be reached when the ratio of landslides in instable zone (high and very high risk) to the relative area of the zone increases. Inversely, the lower ratio of landslides in stable zone (low and very low risk) to relative area of this zone indicates higher overall accuracy. Based on this technique, the values of overall accuracy for the susceptibility map of 1st to 4th approaches were computed as 2.01, 1.94, 0.54, and 1.03 respectively. The magnitudes of the numbers indicate their accuracy. These values convinced us to choose WOE-ANN approach as the most accurate one. The use of pairwise training network of effective factors and extraction of total weights from the AHP comparison matrix in WOE-ANN-AHP approach was a new experiment. But the experiment did not increase the accuracy, since the complexity is not considered when all the layers taken in a network simultaneously. It should be pointed out that the GLR did not suit this study. The reason may be attributed to the nonlinearity of the problem. In addition it appears that the use of UCUs in the UCU-ANN and UCU-GLR approach did not have an effective role in increasing the accuracy. Therefore, it can be concluded that ANN is more flexible method but need more information and is more time consuming than statistical methods.

References:

[1] Nefeslioglu, H.A., Duman, T.Y., Durmaz, S., 2008. Landslide susceptibility mapping for a part of tectonic Kelkit Valley (Eastern Black Sea region of Turkey). Geomorphology, 94 (3-4), 401-418.

[2] Neaupane, K.M., Piantanakulchai, M., 2006. Analytic network process model for landslide hazard zonation. Engineering Geology. 85 (3-4), 281-294.

[3] Varnes, D.J., 1984. Landslide Hazard Zonation: A Review of Principles and Practice. UNESCO Press, Paris, P. 176.

[4] Van Westen, C.J., Van Asch, Th.W.J., Soeters, R., 2005. Landslide hazard and risk zonation: why is it still so difficult? Bulletin of Engineering Geology and the Environment. 65 (2), 176-184.

[5] Soriso Valvo, M., 2002. Landslides; from inventory to risk. In: Rybár, Stemberk, Wagner (eds). Proc. of the 1st European conference on landslides, Prague. Balkema, Rotterdam, 59–79.

[6] Dai, F.C., Lee, C.F., Ngai, Y.Y., 2002. Landslide risk assessment and management: an overview. Engineering Geology. 64 (1), 65–87.

[7] Crozier, M.J., Glade, T., 2005. Landslide hazard and risk: issues, concepts and approach. In: Glade, T., Anderson, M., Crozier, M.J. (Eds.), Landslide Hazard and Risk. Wiley, Chichester, 1–40.

[8] Soeters, R., van Westen, C.J., 1996. Slope instability recognition, analysis, and zonation. In: Turner, A.K., Schuster, R.L. (Eds.), Landslides: Investigation and Mitigation. Transportation Research Board Special Report 247, 129–177.

[9] Yalcin, A., 2008. GIS-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in Ardesen (Turkey):Comparisons of results and confirmations. Catena 72, 1–12.

[10] Ayalew, L., Yamagishi, H., 2005. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. Geomorphology. 65 (1-2), 15-31.

[11] Carrara, A., Guzzetti, F., editors 1995. Geographical Information Systems in Assessing Natural Hazards. Kluwer Academic Publisher, Dordrecht, The Netherlands.

[12] Barredol, J.I., Benavidesz, A., Herhl, J., van Westen, C.J., 2000. Comparing heuristic landslide hazard assessment techniques using GIS in the Tirajana basin, Gran Canaria Island, Spain. International Journal of Applied Earth Observation and Geoinformation. 2 (1), 9-23.

[13] Su¨zen, M.L., Doyuran, V., 2004. Data driven bivariate landslide susceptibility assessment using geographical information systems: a method and application to Asarsuyu catchment, Turkey. Engineering Geology. 71 (3-4), 303–321.

[14] 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.

[15] Komac, M., 2006. A landslide susceptibility model using the Analytical Hierarchy Process method and multivariate statistics in perialpine Slovenia. Geomorphology. 74 (1-4), 17–28.

[16] Saaty, T.L., 1999. Fundamentals of the analytic network process. Proc of International Symposium on Analytical Hierarchy Process, Kobe, Japan.

[17] 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.

[18] Ercanoglu, M., Gokceoglu, C., 2004. Use of fuzzy relations to produce landslide susceptibility map of a landslide prone area (West Black Sea Region, Turkey). Engineering Geology. 75 (3-4), 229–250.

[19] Lin, M.-L., Tung, C.-C., 2003. A GIS-based Potential Analysis of the Landslides

induced by the Chi-Chi Earthquake. Engineering Geology. 71 (1-2), 63–77.

[20] Thiery, Y., Malet, J.P., Sterlacchini, S., Puissant, A., Maquaire, O., 2007. Landslide susceptibility assessment by bivariate methods at large scales: Application to a complex mountainous environment. Geomorphology. 92 (1-2), 38–59.

[21] Carrara, A., Crosta, G., Frattini, P., 2003. Geomorphological and historical data in assessing landslide hazard. Earth Surface Processes and Landforms. 28 (10), 1125–1142.

[22] Dai, F.C., Lee, C.F., 2003. A spatiotemporal probabilistic modeling of storm-induced shallow landsliding using aerial photographs and logistic regression. Earth Surface Processes and Landforms. 28 (5), 527– 545.

[23] Yesilnacar, E., Topal, T., 2005. Landslide susceptibility mapping: A comparison of logistic regression and neural networks methods in a medium scale study, Hendek region, Turkey. Engineering Geology. 79 (3-4), 251–266.

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

[25] Ermini, L., Catani, F., Casagli, N., 2005. Artificial Neural Networks applied to landslide susceptibility assessment. Geomorphology. 66 (1-4), 327–343.

[26] Neuhäuser, B., Terhorst, B., 2007. Landslide susceptibility assessment using weights-of-evidence applied to a study area at the Jurassic escarpment (SW-Germany). Geomorphology. 86 (1-2), 12–24.

[27] Calcerrada, R., Luque, S., 2006. Habitat Quality Assessment using Weight-of-Evidence based GIS modeling: The Case of Picoids Tridactylus as Species Indicator of the Biodiversity Value of the Finnish Forest. Ecological Modeling, 196 (1-2), 62-76.

[28] Vahidnia, M.H., Alesheikh, A.A., Alimohammadi, A., Hosseinali, F., 2009. Hospital site selection using fuzzy AHP and its derivatives. Journal of Environmental Management. 90(10), 3048-3056.

[29] Boroushaki, S., Malczewski, J., 2008. Implementing an extension of the analytical hierarchy process using ordered weighted averaging operators with fuzzy quantifiers in ArcGIS. Computers & Geosciences. 34 (4), 399–410.

[30] Malczewski, J., 1999. GIS and Multicriteria Decision Analysis. Wiley & Sons INC.

[31] Graupe, D., 2007. Principles of Artificial Neural Networks. World Scientific Publishing.

[32] Beale, R., Jackson, T., 1998. Neural Computation: An Introduction. Institute of Physics Publishing.

[33] Paola, J.D., Schowengerdt, R.A., 1995. A review and analysis of backpropagation neural networks for classification of remotely sensed multi-spectral imagery. International Journal of Remote Sensing. 16 (16), 3033-3058.

[34] Picton, P., 2000. Neural Networks. Palgrave Macmillan.

[35] Vining G., 1997. Statistical Methods for Engineers (Hardcover). Duxbury Press.