GIS STATISTICAL SHALLOW LANDSLIDE ASSESSMENT AT A SUB-CATCHMENT SCALE IN NORTH-WESTERN VIETNAM

Fabiana Viscarra Agreda[1]

Abstract

Two small sub-catchments of Lai Châu province in the north-western part of Vietnam are under study due to unceasing landsliding activity. For such purpose, landsliding factors are combined, assessed and ranked with statistical methods; weathering processes, geomorphological complexes, fault density, distance to road, rainfall patterns and slope gradients are determined as most causative factors among others like topographical elevation, geology, distance to rivers, and vegetation. Based on 50 monitored shallow landslides in 2006, the obtained susceptibility map with the statistical weighting method is selected as the most accurate, since it correctly predicted 86% of the landslides.

Key words: slope stability; shallow landslide; statistical methods; landslide susceptibility zonation.

[1] MSc. Licenciado en Ingeniería Civil – Universidad Privada Boliviana, fviscarra@gmail.com

1 INTRODUCTION

Unceasing landsliding activity has been reported in the north-western part of Vietnam, the presence of a Monsoon climate combined with mountainous hilly-topography set the suitable conditions for shallow mass movements triggered by heavy rainfall. For the validation of such theory, two small sub-catchments of Lai Châu province are under the scope of the present study. Data collection and preparation was performed by the Vietnamese Institute of Geo-science and Mineral Resources VIGMR for the assessment of a 2.132 km^2 project, which includes the two sub-catchments under study.

Landslide susceptibility is defined as the terrain proneness to produce slope failure and is usually expressed in a cartographic manner (Brabb, 1984). A landslide susceptibility zonation LSZ map portrays those likely sliding areas and they are normally obtained by combination of causative factors.

Landslide hazard maps are frequently used in regional and urban planning; they may indicate where movements may occur but rarely the circumstances under which they were driven. Terlien (1998) accredited that landside-triggering mechanisms should be included in hazard maps as complementary information, in order to understand better the nature of landslide, to keep a historical record and determine a suitable stabilization procedure.

Some essential principles should be considered at time of landslide susceptibility zonation:

- Slope failures leave distinguishable features that can be recognized through aero-photogrammetry and remote sensing analysis.
- Landslides are likely to occur under similar geologic, geomorphic and hydrologic conditions of past events.
- Settling conditions for landsliding can be direct or indirectly related to future landslide occurrence; since landslide movements are controlled by mechanical laws, they can be simulated through empirical, statistical or deterministic evaluations.
- Landslide occurrence in space or time can be estimated through heuristic investigations or from physical based assessments.

GIS modeling methods have been widely and successfully employed for ranking causative factors and determining their susceptibility landsliding. They can be distinguished as: qualitative or quantitative, and direct or indirect (Van Westen, 1993; Carrara, 1983; Soeters and Van Westen, 1996). Table 1 groups most of the common methodologies into five categories: direct geomorphological mapping, analysis of landslide inventories, heuristic or index based methods, statistical methods and process based conceptual.

Table 1 - Different types of landslide susceptibility methods and their characteristics (Van Westen 1997)

	Direct	Indirect	Qualitative	Quantitative
Geomorphological mapping	✓		✓	
Heuristic (index-based)		✓	✓	
Analysis of inventories		✓		✓
Statistical modeling		✓		✓
Process based (conceptual)		✓		✓

Qualitative methods are subjective and discern different susceptibility levels through descriptive terms. Whilst quantitative methods use numerical based estimations to differentiate susceptibility zones and evaluate landslide hazard.

Geomorphological mapping of landslide susceptibility is a direct mapping method that identifies spatial distribution of unstable locations based on geomorphological characteristics of existing landslides.

Indirect methods estimate potential instability through the analysis of causative landsliding factors. They follow chronological procedures, for instance recognition and mapping of landslides over a target region, has to be done as first attempt followed by the identification and mapping of physical factors directly or indirectly related to landsliding.

The present work applies two statistical methods in order to evaluate the importance of main influencing factors and perform a suitable hazard zonation for a tropical abrupt region in northwestern Vietnam.

2 STUDY AREA

Ancillary data used for the present research is part of a 2,132 km2 risk assessment data-collection project in Tây Bắc. Bounding coordinates for the entire Vietnamese project are 285,470; 2,406,700 and 324,410; 2,461,450 located at zone 48 according to UTM projection with geoid WGS84. Figure 1 displays the Vietnamese region's location and the area studied by VIGMR in 2006, where the two small-enclosed sub-catchments (Nam Lay and Le Bau) were taken into account for the present research.

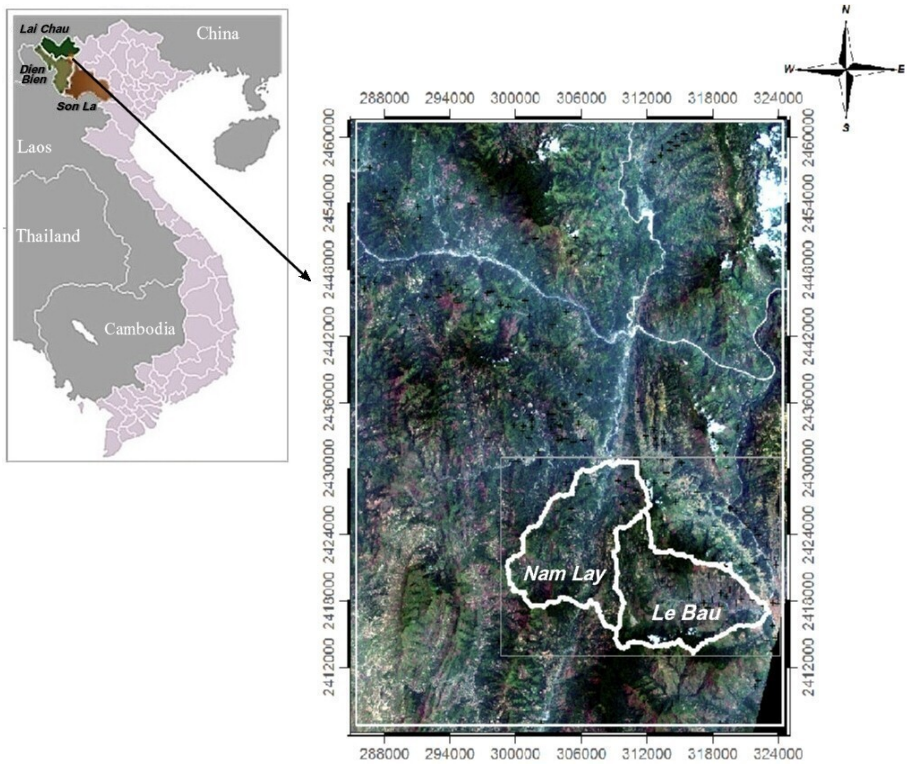

Figure 1 – The two sub-catchments of study and the 2.132 km2 land-slide assessment project of the VIGMR in Lai Châu and Son La provinces (north-west Vietnam).

2.1 Landslide inventory

Both small sub-catchments were considered into a 1:50.000 scale study performed in 2006 by VIGMR. 225 shallow landslides were selected from field recognition within a 2.132 km2 Vietnamese project, considering mostly slides initiated at hill tops or gully heads that have advanced and eroded their path after a presumably heavy rainfall event. Table 2 describes the frequency of landslide events related to surface areas under consideration, where both sub-catchments have a total area of 207.7 km2 with 50 observed landslides.

Table 2 - Data facts about monitored regions with number of identified landslides

Region	*Area (km2)*	*# LS contained*
Whole project	2131.97	225
Nam Lay sub-catchment	108.07	21
Le Bau sub-catchment	99.65	29

3 METODOLOGY

The present work develops two bivariate statistical approaches for the assessment of main factors influencing shallow landsliding. Landsliding factors were quantified, combined, evaluated and ranked in detail by means of these methods.

The methodology mainly focuses on four stages:

3.1 Ancillary data collection

Available topographical maps, geological maps, geomorphological maps, land-use maps (at approximate scale 1;50,000) were collected, digitalized, georeferenced, corrected and geo-rectified in order to generate polygon and raster images for GIS assessment purposes. Landsat ETM+ images were used for land-use evaluation and 10 year average monthly precipitation data from the closest station (Lai Chau) were the main input data throughout the research.

3.2 Data Preparation

Despite geological, geomorphological, weathering and land-use maps, other additional maps were prepared to perform the statistical methods and to determine the degree of influence for each possible factor and their sub-classes.

Elevation values from created Digital Elevation Model DEM ranged from 220 to 1960 meters over sea level, and some derivative maps (i.e. Elevation, Slope and Aspect) were satisfactory created from the DEM. Other built maps based upon available ancillary data like: density fault map, drainage density map, distance to rivers and streams map and distance to roads map, were also created.

Figure 2 shows the scheme pursued in order to prepare the data and generate different landsliding causative factor maps; whereas Table 3 displays the importance of such factors at time of landslide zonation assessment whether for statistical or deterministic methods.

Table 3 – The different factors and their relevance at time of Landslide Hazard Zonation for statistic and deterministic methodologies

Data Layer and types	Relevance for LHZ	Map:	STATISTICAL	DETERMINISTIC
DEM		elevation	✓	
Slope gradient	is the most important factor in gravitational movements	slope	✓	✓
Slope direction	might reflect deifferences in soil moisture and vegetation that can influence stability		✓	
Slope shape	concavity/convexity are indicators for slope hydrology		✓	
Contributing areas	indicator for wetness of slope	(a/b)	✓	✓
Drainage density	might be a good indicator for the type of terrain	river_density	✓	
GEOLOGY	(mainly for rockslides: lithological map, discontinuities, geol struct. Related slopeand d	geology	✓	
Weathering	Depth of weathering profile is important for landslide ocurrence	Weathering	✓	✓
Faults	Distance from faults is important for predictive mapping	Fault_density	✓	
SOILS				
Soil types	Engineering soil types related to geotechnical classification	soil_type		✓
Soil depth	crucial for stability analysis	soil_depth		✓
Geotechnical properties	Cohesion, friction angle, dry unit weigth, saturated unit weigth, porosity	c, f, γdry, γsat, Gs, η		✓
Hydrological properties	Saturated conductivity, field capacity, degree of saturation Sr	K, f, θfc		✓
HYDROLOGY				
Water table	Spatially and temporal varying depth to ground water table	$m_{(0)}$, m_{crit}		✓
Hydrologic component	Long-term annual total precipitation, from 10 year database	rainfall	✓	
	Intensity-Duration-Frequency from 21 year storm database			✓
Stream network	Buffer zones around first, second order streams, or around eroding rivers	dist_to_stream	✓	
GEOMORPHOLOGY				
Geomorphological units	gives small subdivision of zones, relevant for mapping	geomorphology	✓	
LANDUSE				
Vegetation	Vegetation type, canopy cover, root cohesion, surcharge.	landuse	✓	✓
Roads	Buffers around roads and highways	dist_road	✓	

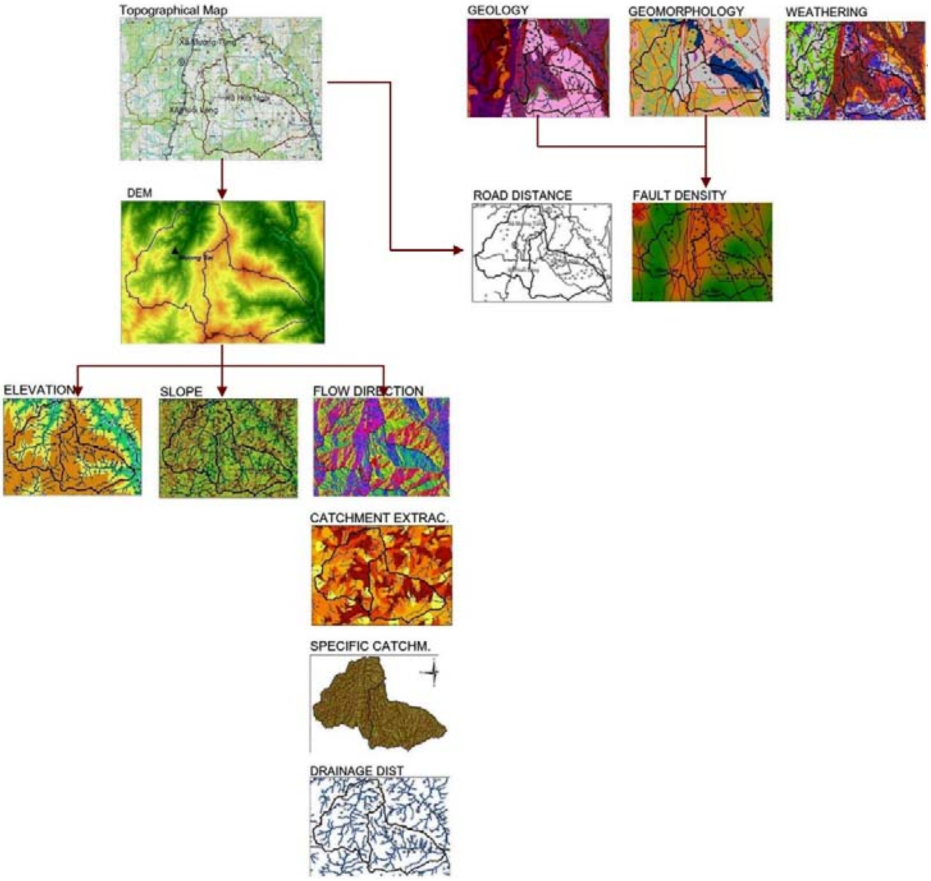

Figure 2 – Scheme of data preparation from the collected ancilliary data.

3.3 GIS statistical analysis

Statistical Bivariate Landslide Susceptibility models quantify the importance of factors leading to mass movement through density calculation of landslides within each causative-factor or parameter map. And within each parameter map's class in order to get a combined map of driven weights.

Statistical index method (W_{ij})

The statistical index method assigns a certain weight value Wij to each parameter class e.g. a sialferite SFA weathering crust or a specific fault density class. Van Westen (1997) defined the Wij value as the natural logarithm of landslide density within the class divided by general landslide density in the entire map:

$$W_{ij} = Ln\left(\frac{f_{ij}}{f}\right) = Ln\left(\frac{N_{ij}}{A_{ij}} \cdot \frac{A_T}{N_T}\right) \tag{1}$$

where fij is the landslide density within class i of parameter j (Nij/Aij); f is total landslide density within the entire study area (NT/AT); Nij is number of landslides within class i of parameter j; Aij is the area of class i enclosed in parameter j; NT is total number of landslides within the entire map, and AT is total area of the entire map.

Statistical correlation between landslide inventory map and attributes from different available parameters is the main basis of the statistical index Wij. Ranges of Wij values give valuable information about the incidence of certain parameter to landslide occurring. Higher Wij values categorize certain class as potential zones for mass movement initiation.

Statistical weighting factor (Wj) method

The statistical index method attributes weighting values Wij=1, assuming that all parameter maps present an equal instability effect, which is unrealistic. Hence, an overall weighting parameter factor Wj was used by Cevic et al. (2003), Oztekin et al. (2005), expressed by:

$$W_j = \frac{(TW_{ij}) - (Min\_TW_{ij})}{(Max\_TW_{ij}) - (Min\_TW_{ij})} \tag{2}$$

where Wj is the weighting factor for each j parameter map; TWij is the total statistical index value obtained from all pixels that belong to parameter j; Min_TWij is the minimum value of the total statistical index map Wij, , and Max_TWij is the maximum value of the total statistical index map Wij.

Equation 2 performs a stretching by using maximum and minimum values of the Wi;, map. Once Wj values were calculated for each parameter map, the new landslide susceptibility index map was accomplished performing a weighted linear sum (Voogd, 1983).

$$LSI = \sum_{j=1}^{n} W_j \cdot W_{ij} \tag{3}$$

where LSI is landslide susceptibility index; Wj is the weighting factor of j parameter; Wij is value obtained through the statistical index method for certain class i of parameter j, and Wj values might vary from 0 to 1.

Even though, the determination of most causative parameters is normally done by a field expert, it is worth to remark that such statistical information is really valuable to compare the sliding incidence among the selected parameters. Further application for heuristic methods can be guided with Wj obtained values.

Figure 3 resumes the steps to obtain Landslide Zonation Maps for both methodologies: the Statistical Index Method and the Statistical Weighting Factor Method. Weighting values for each sub-class factor are calculated with Eq.1; new weighting maps are generated and combined in order to calculate the LSI value (Eq. 3). LSZ maps are obtained after a frequency classification to determine the ultimate Low, Medium, High and Very High Landslide susceptibility classes.

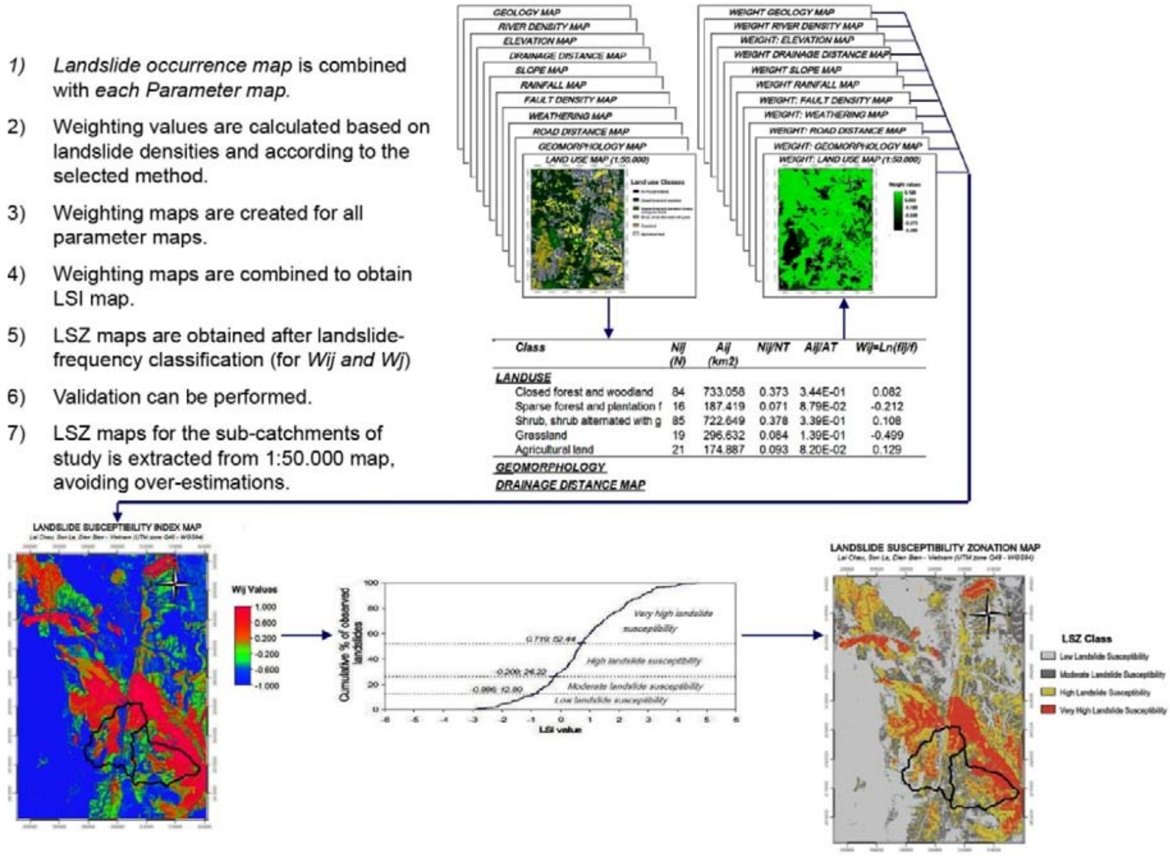

Figure 3 – Scheme of the Statistical Bivariate methodologies.

For example, Figure 4 displays the bordering selected values from the landslides frequency curve. Such graph shows cumulative pixels frequency in percentage versus LSI values after performing a cross function between LSI map and landslide pixels. Limit borders of susceptibility classes were set up in a way that a 1:2 ratio was kept between the susceptibility classes.

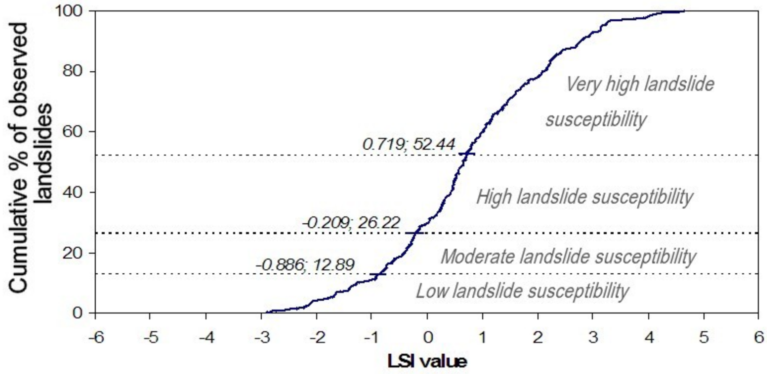

Figure 4 - Cumulative percentage versus LSI values of observed landslides, obtained with statistical weighting factor W_j method, for 225 monitored landslides.

3.4 Validation

In order to assess resulting landslide susceptibility zonation maps validation processes were performed by making a contrast between the obtained LSZ classes and the landslide inventory map. For such purpose, 75 % percent of monitored landslides (training data) were randomly selected and correct landslide prediction accuracy was compared against the 25% left of the inventory population (target data).

4 RESULTS

The bivariate weighting statistical Wj method was selected as the most accurate from the two assessed statistical methods. Although Wij and Wj methods presented a high correlation coefficient (r= 0.75), the selection of Wj as the most efficient methodology is based on the accurate prediction of the observed landslides. Figure 5 displays the susceptibility zonation map of the Wj method for the Nam Bay and Le Bau sub-catchments.

Table 4 resumes the percentage areas for the different landslide susceptibility classes obtained with the studied bivariate statistical methods. Whereas, Table 5 displays the percentage of matched landslides for the different susceptibility classes.

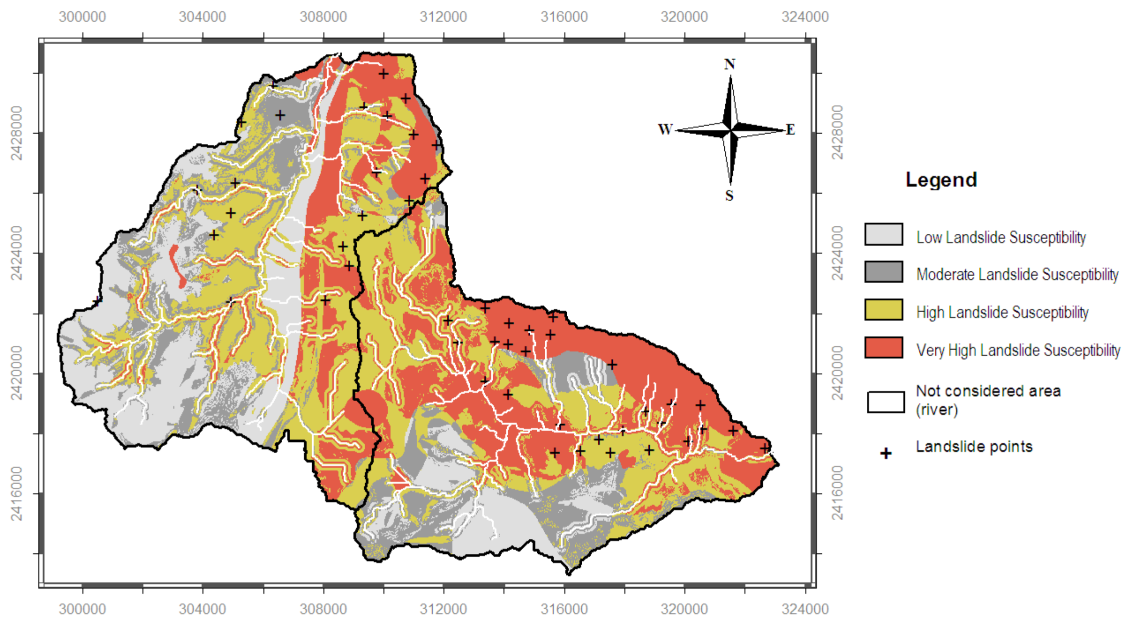

Figure 5 - Landslide susceptibility zonation map for Nam Bay Le Bau ub-catchments with statistical weighting *Wj* method.

It is worth to remark an important accuracy for both statistical methods. For instance 56.0% of the 50 monitored landslides within the studied sub-catchments are well predicted by the Wj method since they were located within the very high landslide susceptible zone, in contrast with 54 % of the Wij method.

Table 4 - Comparison of resulting areas for the different susceptibility zonation classes, obtained with different scenarios and methodologies (Nam Bay and Le Bau sub-catchments).

Landslide susceptibility	***Area of landslide susceptibility class (%)***	
	Bivariate Wij	*Bivariate Wj*
Very high	26.4	29.3
High	33.6	34.8
Moderate	24.0	17.0
Low	16.0	18.9

Table 5 - Comparison of observed landslides (%), that fall within the susceptibility zonation classes obtained with the different methodologies.

Landslide susceptibility	***Observed landslides matching landslide susceptibility class (%)***	
	Bivariate Wij	*Bivariate Wj*
Very high	54.0	56.0
High	34.0	30.0
Moderate	10.0	12.0
Low	2.0	2.0

In this way, an overall accuracy can be evaluated based on landslides observations and by grouping those landslides pixels that were correctly forecasted as very-high or high susceptibility class. Table 6 displays the number and percentage of observed landslides that were correctly or wrongly predicted after grouping the landslides susceptibility zones. The bivariate statistical weighting method predicted correctly 43 of the 50 mass movements (86.0%).

Table 6 - Overall accuracy of landslide prediction, for the statistical weighting method *Wj*.

Accuracy of prediction	***Observed landslides***	
	Number	***percentage%***
Statistical weighting method		
Well (very high + high susceptibility)	43	86.0
Wrong (moderate and low susceptibility)	7	14.0

It is clear how both statistical methods presented a high accuracy (> 70%) for well predicted landslide occurrences. Landslides were matched as "good" if fell into high or very high landslide categories

Table 7 displays the ranked parameters obtained with the Wj method (from most causative to less landsliding causative factors). Geomorphology turned to be the most decisive parameter, having an weighting factor of 63.6%; in the same way sliding due proximity to roads (54.8%), sort of weathering crust (44.9%), fault density (40.3%), distribution of long-term annual precipitation (38.4%) and slope gradient (35.8%) should definitely be considered as the most influencing factors for the landslide assessment.

Finally, landuse (33.1%), distance to stream and rivers (33.2%), elevation of the terrain (32.2%) presented a lower repercussion, leaving behind geology as probable least influencing parameter with a weighting value of 6.4%.

Table 7 - Weighting values for the 10 landslide causative factors (where Min_TWij had a value of -2.888 and Min_TWij was equal to 4.656).

Parameter map	***TWij***	***Wj***
Geomorphology	1.911	0.636
Road Distance	1.248	0.548
Weathering	0.50	0.449
Fault density	0.150	0.403
Rainfall	0.005	0.384
Slope	-0.187	0.358
Landuse	-0.392	0.331
Drainage distance	-0.384	0.332
Elevation	-0.460	0.322
River density	-0.646	0.297
Geology	-2.407	0.064

Table 8 - Accuracy of LSI new categorization with 75% of total monitored landslides (training data) and 25% (target data). *Wrong* accuracy is obtained by summing low and moderate landslide, whereas *good* accuracy by summing: very high and high susceptibility.

Accuracy of prediction	***Training data: 75% Observed landslides***		***Target data: 25% Observed landslides***		***Overall data: 100% Observed landslides***	
	Number	***percentage%***	***Number***	***percentage%***	***Number***	***percentage%***
Statistical index method						
Wrong	59	26.22	57	25.33	59	26.22
Good	166	73.78	168	74.67	166	73.78
Statistical weighting method						
Wrong	55	24.44	66	29.33	59	26.22
Good	170	75.56	159	70.67	166	73.78

From the validation process, correctly and wrongly predicted landslide percentages for training and target data resulted quite similar. LSZ maps obtained with both methods presented reliable results for sliding susceptibility. Both final maps obtained for statistical and weighting methods presented the same amount of low plus moderate landslide (59 pixels) and for high plus very high classes (166 matched pixels) for the whole VIGMR area project, as shown in Table 8.

5 CONCLUSIONS

The selection criteria for the most accurate LSZ map is based on forecasting precision for high and very-high landslide susceptibility zones. Results showed that the *Wj* method yields the best results for landslide susceptibility mapping, since 43 of the 50 observed landslides were correctly predicted (86.0%); whilst, only 7 landslides were wrongly predicted (14.0%).

Thereby, the statistical weighting factor method *Wj* is the approach that yields the best shallow landslide susceptibility map for Nam Bay and Le Bau sub-catchments. It is also important to remark that statistical methods are not capable to predict those regions highly affected by a sudden water table increase as deterministic methods do. Moreover, statistical methods are greatly influenced by observed landslides subjected to high error risks from possible wrong performed monitoring.

From the statistical *Wj* factor methodology, it can be stated that weathering processes, geomorphological complexes distribution, fault density, closeness to roads, rainfall patterns and slope gradients were the most causative factors for shallow landsliding. In addition, other factors like topographical elevation, geology, distance to stream and rivers, and vegetation, were also evaluated, but did not present considerable instability effects.

6 REFERENCES

BRABB, E. (1984). Innovative approaches to landslide hazard and risk mapping. In: IV INTERNATIONAL SYMPOSIUM OF LANDSLIDES, Vol. 1, Toronto, Canada: p.307-323.

CARRARA, A. (1983). Multivariate Models For Landslide Hazard Evaluation. **Mathematical Geology**, 15(3): p. 403-426.

CEVIK, E., and TOPAL, T. (2003). GIS-Based landslide susceptibility mapping for a problematic segment of the natural gas pipeline, Hendek (Turkey). **Environmental Geology**, 44: p. 949-962.

OZTEKIN, B. and TOPAL, T. (2005). GIS-based detachment susceptibility analysis of a cut slope in limestone, Ankara-Turkey. **Environmental Geology**, 49, p. 124-132.

SOETERS, R. and VAN WESTEN, C. (1996). Slope instability recognition analysis and zonation. In: TURNER KT, SCHUSTER RL (ED.) LANDSLIDE: Investigation and Mitigation. Spec. Report 47, Transportation Research Board, National Research Council, Washington D.C., pp. 129-177.

TERLIEN, M. T., 1998. The determination of statistical and deterministic hydrological landslide-triggering thresholds. **Environmental Geology** 35: 124-130.

VAN WESTEN, C., 1993. Application of geographic information system to landslide hazard zonation. INTERNATIONAL INSTITUTE FOR AEROSPACE AND EARTH RESOURCES SURVEY, Enshede, The Netherlands, ITC-Publication N. 15: 245 p.

VAN WESTEN, C., 1997. Statistical landslide hazard analysis. ILWIS 2.1 for windows application guide. Enshede, The Netherlands, ITC Publication N. 15: 73-84.