Exploratory GIS & remote sensing analysis for the development of statistical correlations between environmental parameters and mass movements' occurrence

Chadi Abdallah

National Council for Scientific Research/Remote Sensing Center, P.O. Box 11-8281, Beirut, Lebanon;
Tel: +96-4-409845/6; Fax: +961-4-409847; Email: chadi@cnrs.edu.lb

1. Abstract

Among the various natural hazards, mass movements (MM) are probably the most damaging to the natural and human environment in Mediterranean countries, including Lebanon which represents a good case study. This research deals with how to use Geographic Information Systems (GIS) for establishing the relationships between MM occurrence and different factor terrain parameters over a representative region of Lebanon. Parameters expressed by: 1—ancillary data like lithology, proximity to fault zone, soil type, land cover/use, distance to drainage line and rainfall quantity, and 2—derived data like slope gradient and slope aspect, were correlated with MM using GIS-approaches. MM were detected through visual interpretation of two stereo-pairs of SPOT 4 images (anaglyph) at 10 m resolution. This study indicates, depending on bivariate Remote Sensing and GIS statistical correlations (Kendall Tau-b correlation), that lithology is the most influencing factor on MM occurrence. It also shows that statistical correlations to mass movements exist best between factors at the following decreasing order of importance: lithology–proximity to fault line, lithology–soil type and lithology–distance to drainage line at 1% level of significance, and soil–land cover/use, slope aspect-land cover/use, and soil–slope gradient at 5% level of significance. These correlations were verified and checked through field observations and explained using univariate statistical correlations. Therefore, they could be extrapolated to other Mediterranean countries having similar geoenvironmental conditions.

2. Introduction

Satellite images have been used worldwide to visually identify large landslides without differentiating other types of mass movements (MM) (Zhou and Li, 2000; De la ville et al., 2002; Hervas et al., 2003). Moreover, the reliability in detecting mass movements can differ to a wide extent according to the processing technique used, the sensor chosen and MM dimensions. For instance and at regional level, utilizing Landsat TM (30m) for identifying large MM in vegetated or mountainous areas would not be a very difficult task, since it is uncomplicated to visually discriminated large areas of eroded or washed out terrain cover. But when it comes to discriminate MM of small dimensions at regional scale, in a rugged topographic area (gradient ~ 100-125 m/km) using the formal type resolution of imageries, it seems to be challenging. Abdallah et al. (2007), compared statistically the applicability of different satellite data sensors (Landsat TM, IRS and SPOT), preferred image processing techniques (False Color Composite, pan-sharpen, principal component analysis and anaglyph) and cost for mapping diverse types of MM in rugged topography (Lebanon Mountains). The former study showed that 3D anaglyph SPOT images proved to have the least cost and the best discriminating/sensing the types of MM with a mean visual interpretation error (38%).

Additionally, to sort out parameters influencing the occurrence of MM and determining their relative weights at regional scale is crucial for building susceptibility and hazard MM maps. The susceptibility maps are related to the instability of the terrain itself while the hazard maps take into account also the triggering parameters (either natural or anthropic). Nevertheless, sorting of the parameters that have a role in MM occurrence is not an easy task. There are neither universal criteria nor guidelines (Ayalew and Yamagishi, 2005). The general consensus of sorting the considered parameters must be operational, complete, non-uniform, measurable and non-redundant. However, it is possible to establish a priori list of parameter categories that may intervene. In this present work, six major parameter categories were considered, i.e. geomorphology, geology, soil, hydrology, climate, and anthropic influence responsible for instabilities and triggering conditions, depending on field observations and referring to several works conducted in this domain worldwide. The diagnostic of the relative weights of these parameters can be done using statistical analysis to prevent the subjectivity of the different investigators in assuming the most significant ones depending on their own skill and experience. Up to now, many statistical methods (univariate, bivariate, multivariate correspondence analysis, linear regression, logistic regression) have been used for this determination in many areas of the world (Atkinson and Massari, 1998; Wang et al., 2002). Each method has its advantages and drawbacks (Abdallah, 2007). A binary statistical approach was used in this study to explore dual relationships between terrain parameters according to their importance in causing mass movements. These approaches are used also to determine the most influencing parameters which can be used as weighted input data in prediction susceptibility and hazard MM maps. The Geographic Information System (GIS) was utilized for extracting the influencing parameters as maps needed for studying problems at regional scale, and for overlaying parameter maps and MM map locations obtained through visual interpretation of satellite imageries. This overlaying is essential for building tabular databases on which statistical analysis was run.

In this context, MM were first identified by visual interpretation of 3D anaglyph SPOT4 imageries. The choice and preparation of the considered influencing parameters, their GIS manipulation and statistical running analysis were then exposed. The obtained field, bivariate and univariate statistical correlations between parameters on one side, and between parameters and MM occurrence on the other side were finally calculated.

3. Methods

Anaglyphs are stereo pairs merged together into a single picture. They are constructed by colouring a stereo pair's left side red and the right side blue, or sometimes the other way round. The stereo images are merged so that points in one image approximately coincide with their conjugate points in the other image, but without referencing one to the other. The technique does not depend on parallel line sight, and stereoscopy is more easily attained. Anaglyph images of SPOT-HRV stereo-couples were produced for a representative mountainous area in Lebanon experiencing numerous MM occupying an area of 670 Km^2. The two chosen panchromatic stereo-pair SPOT 4 images are of a high ground resolution (10 m). They were acquired with incident angles of 2.3 and 30.3°, dated on 31 May 2002 and 1 June 2002, respectively. The close dates of image acquisition and the appropriate incident angles between the two SPOT-P, were behind choosing this stereo couple pairs.

For each parameter category (geomorphology, geology, soil, hydrology, climate, anthropic influence), one or several parameters were considered as influencing the occurrence of MM within the study area according to our field observations and the consultation of several works done in similar or dissimilar environmental conditions. In this context, a field format was established for each MM location, taking into account information relative to the lithological structure, soil, vegetation cover, type of land use around the mass movements, slope, elevation, the corresponding causes of MM (road construction, river and water ways, etc.) as well as human activities protecting against this phenomena. The chosen parameters were extracted from remote sensing or ancillary maps.

All the mentioned parameters were handled at different levels of treatment with GIS software (*ArcGIS 9.2*). To obtain data homogeneity, the obtained raster models (elevation, slope gradient, slope aspect, slope curvature and seismic events) were converted from raster to vector using the Spatial Analyst extension of the ArcGIS software. Mass movements areas recognized through visual interpretation of remotely sensed data were delineated and converted "*using ArcGIS appropriate extensions*" into centroides resulting in a point-theme-layer of the distribution of mass movements in the study area. This procedure was undertaken to reduce accuracy errors due to the shifting in overlapping layers and to easily differentiate the detachment areas from the zones of accumulation. Each mass movement point was given an ID number. Overlaying the "centroid" point-theme-layer with every thematic vector terrain parameter has permitted defining the parameter class in which each mass movement falls in each parametric layer. This overlaying is algorithmically simple, fast and easily computed. Thus, resulting attribute tables correlated the occurrence of the detected mass movements with each category characterizing each terrain parameter. These tables were then exported to Microsoft Excel and gathered in a matrix sheet showing in the row fields the seventeen different terrain parameters and in the column fields the numbered mass movements **(Table 1)**.

Univariate statistical correlations were first produced, i.e. relations between mass movement frequency and the different categories of each parameter. On the other hand, non-parametric bivariate procedures of Kendall's tau-b rank correlation coefficients, with their significant levels were computed using the SPSS software package. Both correlations (univariate and bivariate) were verified in the field. In order to determine the statistical correlations between the chosen parameters of different nature, homogenization of their categories is required. Some parameters are quantitative, like elevation, slope gradient, slope aspect, slope curvature, proximity to fault zone, distance to drainage line, distance to water sources, rainfall quantity, proximity to roads and seismic events, while others are qualitative like lithology, soil type, land cover/use, floods and forest fires. This is why the categories of those parameters were qualified by giving a random number for each category in each parameter. The idea was that the chaotic distribution of numbers will not be a weighting factor in the statistical correlations.

4. Results

202 different types of MM was detected using Visual interpretation of Anaglyph imageries. After obtaining the tabular data from the GIS concerned parameters involved in MM, a correlation matrix between the seventeen parameters affecting mass movements was built (Table 1). The analysis of this matrix indicates that the susceptibility of the terrain to mass movements includes all the correlated parameters which are inherent to the terrain itself, but are acting with different weights and levels of significance (1% or 5%). The obtained values vary between -1 and +1, with a positive correlation indicating that the ranks of both parameters increase together, whilst a negative correlation indicates that as the rank of one variable increases, the other one decreases. As a result, some parameters are correlated several times with other parameters at 1% level of significance (major correlation), while others are correlated but at lower level of significance (5% - minor correlation) (Table 1). The number of relationships between the seventeen parameters was equal to 136, among which, 12 correlations are found at high level of significance (1%) and 27 at low level (5%). However, one parameter can be correlated with two different parameters, and those will be also correlated together, all of the three influencing the occurrence of MM. As an example, lithology is highly correlated to the distance to drainage line from one side and to the distance to water courses (springs) from the other side; drainage lines and springs being also highly correlated. Despite of the inherent interdependence of all the physical terrain parameters taken together, histograms were plotted in order to detect the univariate relations between mass movement frequency and the different categories of each parameter. In this context, lithology is the predominant parameter in inducing mass movements, since it shows the highest correlation with other parameters (7 times at 1% level of significance and 3 times at 5%). The correlation is strong (1% significance level) with proximity to faults, karst type, distance to quarries, soil type, distance to drainage line, distance to water sources (springs), and existence of floods, and is of less strength (5%) if slope curvature, proximity to roads and seismic events are considered. Other parameters have an influence on activating MM but with diverse degrees of effect. Among these, two parameters - soil type and distance to water sources (springs) - are correlated 7 times with other parameters (at 1% and 5%). It was expected that soil type can greatly influence the occurrence of mass movements, and this finding was similarly proved in other studies. However, the most important fact considered once testing the GIS bivariate correlations between parameters was the integration of distance to water sources (springs) as a reflection of the effect of groundwater in inducing MM. This effect was not studied in most similar studies conducted at regional scale. It was shown in this work that the outlets of springs may provoke MM under certain conditions. The other considered parameters have certainly an influence on MM occurrence, but this influence was minimal in some cases due to the unavailability of data reflecting more the power of these parameters (e.g., rainfall quantity was used instead of rainfall intensity). These correlations, as well as field observations and the accurate detection of the location of MM are crucial for MM susceptibility and hazard analysis.

5. References

Abdallah, C., 2007. Assessment of remote sensing and geographic information systems for the study of mass movements in Lebanon. Ph.D. dissertation, University of Pierre and Marie Curie Paris 6 (France), 218p + annexes.

Abdallah, C., Chorowicz, J., Bou Kheir, R., Dhont, D., 2007. Comparative use of processed satellite images in remote sensing of mass movements: Lebanon as a case study. International Journal for Remote Sensing, 28(19): 1-20.

Atkinson, P.M., Massari, R., 1998. Generalised linear modelling of susceptibility to landsliding in the Central Apennines, Italy. Computers & Geosciences, 24(4): 373–385

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: 15–31.

De la Ville, N., Chumaceiro Diaz, A., Ramirez, D., 2002. Remote sensing and GIS technologies as tools to support sustainable management of areas devastated by landslides. Environ Dev Sustain, 4: 221–229.

Hervas, J., Barredo, J.I., Rosin, P.L., Pasuto, A., Mantovani, F., Silvano, S., 2003. Monitoring landslides from optical remotely sensed imagery: the case history of Tessina landslide, Italy. Geomorphology, 54: 63-75.

Wang, X., Jiang, Y., Zhao, Y., Pan, L., 2002. Study and application on the visualization of geological information system in landslideYanshilixue Yu Gongcheng Xuebao/Chinese.

Zhou, G., Li, R., 2000. Accuracy evaluation of ground points from IKONOS high-resolution satellite imagery. Photogrammetric Engineering and Remote Sensing, 66(9): 1103-1112.

Table 1 Correlation matrix between the seventeen used parameters affecting mass movements

Coefficient of correlation	ELE	SLG	SLA	SLC	LITH	FAU	KAR	QUA	SOIL	DRA	SPRI	RAIN	LAND	ROAD	SEI	FLO	FIRE
ELE	1	0.011	0.061	0.013	0.001	0	0.121	0.071	0.051	0.005	0.027	0.112	0.009	0.013	0	0.003	0.013
SLG	0.011	1	-0.014	0.027	-0.089	-0.059	0.136	0.119	0.139	-0.082	0.078	-0.057	0.008	0.121	0.013	0.065	0.017
SLA	0.061	-0.014	1	0.007	-0.056	-0.053	0.007	0.015	-0.059	0.029	0.013	0	0.16	0.003	0	0.011	0.127
SLC	0.013	0.027	0.007	1	0.125	0.013	0.002	0.075	0.112	0.128	0.007	0.072	0.055	0.022	0.008	0.067	0.001
LITH	0.001	-0.089	-0.056	0.125	1	0.275	0.268	0.193	0.189	0.179	0.222	-0.001	-0.081	0.111	0.153	0.176	0.002
FAU	0	-0.059	-0.053	0.013	0.275	1	0.033	0.125	0.026	0.079	0.133	0.002	-0.019	0.156	0.256	0.032	0.015
KAR	0.121	0.136	0.007	0.002	0.268	0.033	1	0.005	0.135	-0.156	0.227	0.003	0.021	0.032	0.05	0.051	0.035
QUA	0.071	0.119	0.015	0.075	0.193	0.125	0.005	1	0.101	0.009	0.156	0.022	-0.156	0.07	0.169	0.003	0.065
SOIL	0.051	0.139	-0.059	0.112	0.189	0.026	0.135	0.101	1	0.033	0.145	0.058	0.167	0.031	0.033	0.116	0.007
DRA	0.005	-0.082	0.029	0.128	0.179	0.079	-0.156	0.009	0.033	1	0.177	0.105	0.063	0.068	-0.013	0.212	0.002
SPRI	0.027	0.078	0.013	0.007	0.222	0.133	0.227	0.156	0.145	0.177	1	0.116	0.007	0.013	0.002	-0.037	-0.017
RAIN	0.112	-0.057	0	0.072	-0.001	0.002	0.003	0.022	0.058	0.105	0.116	1	0.005	0.002	0.015	0.079	0.088
LAND	0.009	0.008	0.16	0.055	-0.081	-0.019	0.021	-0.156	0.167	0.063	0.007	0.005	1	0.176	0.013	0.135	0.111
ROAD	0.013	0.121	0.003	0.022	0.111	0.156	0.032	0.07	0.031	0.068	0.013	0.002	0.176	1	0.001	0.013	0.015
SEI	0	0.013	0	0.008	0.153	0.256	0.05	0.169	0.033	-0.013	0.002	0.015	0.013	0.001	1	0.091	0.003
FLO	0.003	0.065	0.011	0.067	0.176	0.032	0.051	0.003	0.116	0.212	-0.037	0.079	0.135	0.013	0.091	1	0.001
FIRE	0.013	0.017	0.127	0.001	0.002	0.015	0.035	0.065	0.007	0.002	-0.017	0.088	0.111	0.015	0.003	0.001	1

ELE = elevation; SLG = slope gradient; SLA = slope aspect; SLC = slope curvature; LITH = Lithology; FAU = Proximity to fault zone; KAR = Karst type; QUA = distance to quarries; SOIL = soil type; DRA = distance to drainage line; SPRI = distance to water courses; RAIN = rainfall quantity; LAND = land cover/use; ROAD = distance to roads; SEI = seismic events; FLO = floods; FIRE = forest fires

Correlation is significant at confidence level: 1% (value ≥ |0.76|), 5% (value ≥ |0.101| and < |0.7|