LANDSLIDE HAZARD AND RISK BY GIS-BASED MULTIVARIATE MODELS

F. Guzzetti
CNR-IRPI - via Madonna Alta, 126 - 06128 PERUGIA, Italy

INTRODUCTION

Among natural catastrophes landslides are the most frequent in space and time. The yearly cost of mass-movements world-wide can be estimated in billions of US dollars (Alexander, 1989; Swanston and Schuster, 1989; Schuster and Fleming, 1986; Taylor and Brabb, 1986). Mass movement is caused by the interplay of many factors and therefore several variables, and all their possible combinations, have to be considered when trying to predict landslide occurrence.

Government and research institutions have been working for years to assess this type of hazard and risk, and to seek effective remedial measures (Brabb and Harrod, 1989). Many different methods and techniques have been proposed and tested with various degree of success (van Westen, 1993; Carrara et al., 1992; Hansen, 1984; Brabb, 1984). Late in the seventies, experiments were carried out to test the possible use of statistical multivariate models in assessing landslide hazard in areas of limited extent - few tens of square km - in three areas in Southern Italy (Carrara, 1983).

Results proved that black-box models can successfully predict the distribution of actual and past slope-failures. Despite these encouraging results the approach was not lacking limitations. The most important drawbacks were associated with the great time needed to encode the thematic information, arranged in layers, and in the difficulty in treating efficiently the spatial information used to carry out the statistical analysis.

With the widespread introduction of GIS technology, mostly due to the development of new processors and specialised hardware components - digitisers, scanners, plotters, etc. - and the advances in software packages capable of handling spatial information, some of the limitations could be overcome. Recent experiments of using GIS technology combined with statistical analysis for landslide hazard and risk assessment have proved to be successful and reasonably efficient for areas of limited extent (Carrara et al., 1991; van Westen, 1993).

Despite the, mostly technological, improvements the basic methodology remains the same. A statistical univariate or multivariate - discriminant or regression - model is built from the spatial distribution of past and present failures and of factors controlling the occurrence of mass-movements - namely geology, morphology, land-use, hydrogeology, human impact, rainfall, etc.-. The outcome is then tested analysing residual maps, that is, maps showing model failures. The reliability and extrapolation of the model can also be tested (Carrara et al., 1991).

In the Umbria-Marche region, extending for about 20,000 km^2 in Central Italy, a long-term project aimed at the definition of the regional landslide distribution and the associated hazard and risk was started by CNR-IRPI in the late eighties. Both small-scale and large-scale investigations have been carried out using traditional geomorphological methodologies - namely the interpretation of aerial photographs and field surveys - and new techniques pertaining to the vast realm of the management and analysis of spatial information (e.g., Geographical Information Systems, Digital Elevation Models, Statistical modelling, etc.) (Laurini and Thompson, 1992; Aronoff, 1989; Burrough, 1986).

LARGE SCALE CASE STUDIES

Four test areas were selected to carry out detailed, large scale experiments on the feasibility of assessing landslide hazard and risk through GIS techniques and multivariate statistical modeling. In the followings we report on two of these test sites: the Tescio (59 km^2) and Carpina (67 km^2) basins located only 40 km apart and pertaining to the same structural, geological and physiographic domain.
The data needed for the investigation (Table 1) were derived from or obtained by existing topographic maps, aerial photographs, and field surveys. Contour lines at 1:25,000 scale were acquired in digital format from the Italian Military Cartographic Institute (IGMI). After error checking and validation, contours were processed to generate a high-fidelity digital elevation model (DEM) for each basin. The DEMs were further processed through specialised software packages, (Carrara A., 1988) to automatically identify drainage lines, thus subdividing the terrain into morphological entities called *slope-units*.

	TESCIO number of continuous or categorical variables	**CARPINA** number of continuous or categorical variables
Morphological Attributes from DTM	21	25
Lithology	10	9
Bedding Attitude	3	7
Structural Domains	-	12
Hydrogeological Setting	5	6
Land-use	10	9
Landslide	10	11

Table 1. Thematic information mapped for the Tescio and Carpina basins. The original number of classes is given. Some classes were aggregated for the statistical analysis.

As compared to more traditional and widely used terrain units, such as grid-cells, homogeneous domains and land-units, these spatial entities, representing the statistical samples used for modelling purposes, have conceptual and practical advantages and limitations. They reflect meaningful morphological subdivisions of terrain (i.e., main slopes) that relate to the geomorphic process under investigation (i.e., landslides); and they help in reducing mapping errors, particularly those associated with landslide identification (Carrara et al., 1992). Limitations are due to the complexity of the algorithms capable of determining drainage and divide lines in all circumstances, and in the definition - and computation - of morphometric attributes, such as slope length or slope profile form - considered important for landslide hazard and risk assessment.

Geological and lithological data were obtained by mapping the territory at 1:10,000 scale. A large number of structural measurements were taken as uniformly as possible throughout the study areas. In order to assess the hydrogeological conditions of slopes, simple parameters were estimated in the field and from the geological map, or automatically obtained from the elevation model, namely bedding attitude, slope aspect, and the stratigraphic relations between permeable and impermeable rock units. Land use data were derived from aerial photographs with minor field checking. Landslide distribution and classification (Varnes, 1978) - according to movement topology, relative age, degree of activity, estimated depth and velocity, and type of material) - were determined both by the interpretation of aerial photographs and by systematic field surveys.

Table 2 summarises the results of the multivariate - discriminant - statistical analyses carried out for the two basins. The statistical models were correct in predicting landslide prone and not prone areas more than 80% of the time (83.40 % for the Tescio basin and 82.72 for the Carpina basin). This is certainly a good "batting average", possibly better than the success ratio of many surgeons. Both discriminant models were more efficient in predicting stable slopes and less efficient in predicting unstable ones, and, working against safety, incorrectly predicted an excess of stable slope-units.

TESCIO BASIN
266 slope units 15 variables

	PREDICTED	
	STABLE %	UNSTABLE %
STABLE	84.7	15.3
UNSTABLE	18.1	81.9
Percentage of cases correctly classified is 83.40		

CARPINA BASIN
414 slope units 22 variables

	PREDICTED	
	STABLE %	UNSTABLE %
STABLE	83.9	16.1
UNSTABLE	19.7	80.3
Percentage of cases correctly classified is 82.72		

Table 2. Outcome of the discriminant analysis for the Tescio and Carpina basins

Despite the similarities, possibly due to the level of accuracy and quality of the input data, there are substantial differences in the outcomes of the two models. For the Tescio basin (Table 3) some of the variables, namely BOSCO, MAXCA, FORM, IDR_A and IDR_D, have high standardised discriminant function coefficients (SDFC), indicating that these variables were highly successful in discriminating between landslide prone and not prone areas. In the case of the Carpina basin (Table 4) SDFCs have more similar (absolute) values, and no single variable, with possibly the exception of S_BO and FRP_T, is very good in discriminating between the two groups. This may be due to differences in geological, morphological, and land-use setting of the two basins, as well as to the different population size (266 slope-units for Tescio and 414 slope-units for Carpina).

Despite their overall similarity - both basin are located within the same physiographic domain, approximately in the same terrain types, and show similar type of mass-movements - the differences are large enough to produce different statistical outcomes. In the case of the Tescio basin landslide distribution is strongly controlled by the lithologial setting. Mass-movements concentrate in the northern and central part of the basin, where marl and shale units outcrop; and virtually absent in the southern half of the basin, where more competent rocks, chiefly limestone, outcrop. Land-use and hydrogeological conditions of slopes are also different in the northern and southern portions of the basin. This almost binary setting is ideal for the discriminant analysis to give the best results.

	VARIABLE		SDFC
1	CINE	slope-unit % of Scaglia Cinerea unit	-0.202
2	SCHL	slope-unit % of Schlier unit	-0.355
3	ARCA	slope-unit % of sandstone-rich unit	0.331
4	MAXCA	product of marl-rich and calcarenite-rich units	**0.693**
5	DENUD	slope-unit % of uncultivated area	0.314
6	BOSCO	slope-unit % of forest	**-0.601**
7	AN	slope-unit facing N	0.199
8	AW	slope-unit facing W	0.293
9	MAGN	subbasin magnitude	-0.492
10	ELV_M	slope-unit mean elevation	-0.295
11	FORM	slope-unit form - perimeter/area	**-0.503**
12	RXGR	slope-unit surface roughness index	-0.260
13	FRA_TR	bedding dipping toward slope-unit free face	0.251
14	IDR_A	permeable beds capping impermeable ones	**0.545**
15	IDR_D	impermeable beds throughout slope-unit	**0.840**

Table 3 Tescio basin: list of variables entered in the discriminant function and their relative importance as expected by the standardised discriminant function coefficient (SDFC). Low and high values describe respectivelly stable and unstable slopes.

VARIABLE			SDFC
1	NE	slope-unit facing NE	**0.355**
2	SE	slope-unit facing SE	**-0.251**
3	S	slope-unit facing S	**-0.258**
4	ANG_STD	STD of slope angle within slope unit	**-0.268**
5	SLO_LEN	slope-unit lenght	0.213
6	ANGLE3	slope-unit upper third slope angle	-0.246
7	CONC	concave slope-unit profile	0.178
8	CONV	convex slope-unit profile	-0.128
9	COC_COV	concave-convex slope unit profile	-0.092
10	IRR	irregular slope unit profile	0.184
11	CAM	slope-unit % of calcaren., marl, and sandstone unit	**0.359**
12	CALC	slope-unit % of calcarenite unit	-0.165
13	D4	slope-unit % of SW dipping monocline	0.123
14	D6	slope-unit % of NE dipping monocline	-0.136
15	D8	slope-unit % of E-SE dipping monocline	-0.169
16	S_BO	slope-unit % of forest area	**-0.393**
17	S_SAP	slope-unit % of cultivated area	0.176
18	FRP_T	bedding plane toward free face steeper than slope	**0.449**
19	FRM_T	bedding plane toward free face less steep than slope	0.247
20	CATA	fault area within slope-unit	0.248
21	AC_TM	permeable beds capping impermeable ones	0.102
22	AC_A	aquifer in alluvial deposits	-0.245

Table 4. Carpina basin: list of variables entered in the discriminant function and their relative importance as expected by the standardised discriminant function coefficient (SDFC). Low and high values describe respectivelly stable and unstable slopes.

In the case of the Carpina basin, landslides are more scattered throughout the area and a larger set of variables (22 vs 15, Tab. 3 and 4) entered into the discriminant function, among which a few (NE, SE and S) with high SDFC values do not bear a clear physical meaning.

SMALL SCALE STUDY

The Umbria-Marche region has a long history of landslide events. Reports on deep-seated mass-movements and debris flows go back to the 15th century (Guzzetti and Cardinali, 1991; Almagià, 1910). Despite all the research efforts, mostly concentrated on site specific events, or on the identification of landslide types and mechanisms for engineering purposes, the full extent of landsliding in the area has been ascertained only recently with the production of small-scale landslide inventory maps for the whole territory. Maps at 1:100,000 scale were produced through the interpretation of 1:33,000 scale, black and white aerial photographs made during flights in the 1950s. Limited field checks were performed (Guzzetti and Cardinali 1989a, 1989b, Antonini et al., 1993).

Landslides inventory maps are the easiest way to show the spatial distribution of failures (Hansen, 1984). Outlining the areas where landslides have already occurred, they represent the simplest type of landslide hazard maps (van Westen, 1993). The reconnaissance survey carried out in the Umbria-Marche region revealed about 27,000 landslides - the result of possibly more than 10,000 years of geomorphologic evolution of slopes - of different extent, type, age and degree of activity, covering a total of about 1,800 km^2, 9% of the entire area. Due to the reconnaissance techniques used to make the inventory, the scale of the maps and the date of the photographs, mass movement too small to be mapped at 1:100,000 scale was not shown on the maps.

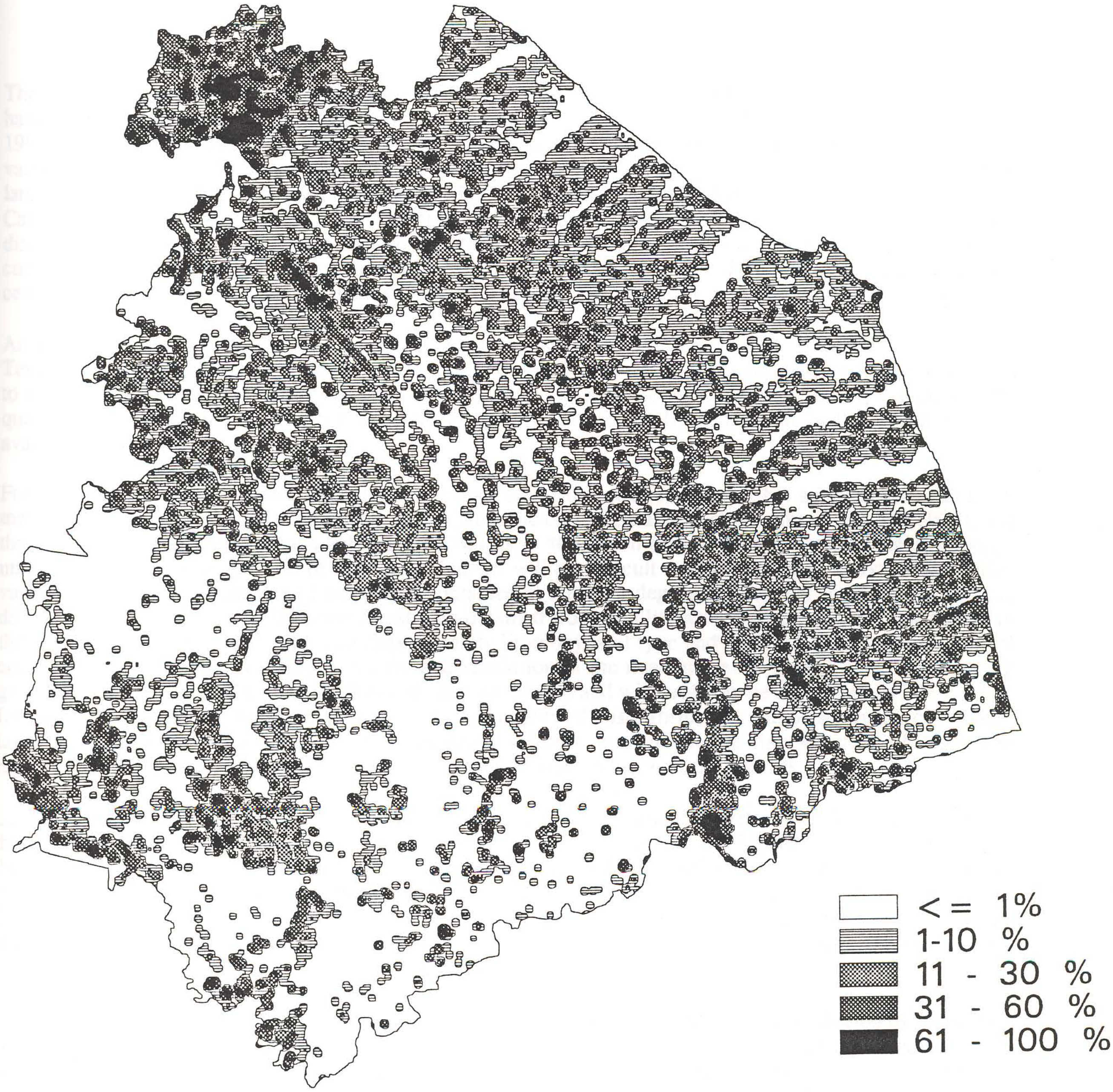

Figure 1. Landslide density map for the Umbria-Marche region

These inventory maps were subsequently digitized. The resulting regional landslide-database was then used to carry out a landslide density analysis. Vector data were rasterized - grid spacing was 100 meters - reduced in binary format - each grid-cell was assigned a landslide/no landslide value - and the percentage of landslide area within a circular moving window of 1 km^2 was computed. The outcome was a landslide isopleth map (DeGraff, 1985) that represents the only small-scale landslide hazard assessment product in the country (Figure 1).

FUTURE DEVELOPMENTS

The advantages and limitations of multivariate statistical analysis supported by GIS technology in landslide hazard and risk assessment, for areas of limited extent, are sufficiently known (Carrara, 1983, Carrara et al., 1991; Carrara et al., 1992; van Westen, 1993). Quality and spatial resolution of data, together with the type of variables and the sampling technique - regular, using grid-cells, or irregular, using aggregate of pixels such as land-units or slope-units - are the crucial issues. The type of statistical tool does not seem to be so important. Carrara et al. (1992) have shown that different multivariate techniques - namely logistic regression and discriminant analysis - give similar (but not identical) results. It remains to be understood how to efficiently combine (different) outcomes from various predictive models, possibly with different levels of accuracy and certainty. Tentatively, the most promising approach to this problem may lay in the use of expert systems.

An open problem is how, and to what extent, the knowledge acquired on large-scale studies (such as the Tescio and Carpina cases) on the *true* spatial distribution of failures and on factors controlling it, can be used to assess regional landslide hazard and risk. It is now clear that in the foreseeable future the type, accuracy and quality of information available for large scale studies, on areas of limited extent, will only partially be available for large regions.

For the Umbria-Marche region, digital contour lines could be acquired from IGMI and detailed - 20 to 25 meter spacing - DEM could be produced. A part from the cost of data acquisition (about 100,000 US $), and the time required to generate such a detailed DTM (about 40 million pixels and an estimated 150,000 slope-units), the efficient management of such large data-set will be difficult even on powerful workstations. Other variables, such as land-use, could possibly be obtained at the right degree of accuracy from remote-sensing data. R. Mark (1992) used a forest/no forest map obtained from satellite data to assess debris flow hazard in the San Mateo County. Still other variables - namely lithology, soil type and thickness, and hydrogeological conditions - will not be available at the desired resolution in the near future. A project aimed at mapping geology for the entire country at 1:50,000 scale has just started and will probably be completed in 20-25 years. It remains unclear if, and to what extent, the type and quality of this information will be adequate for regional landslide hazard and risk assessment. A possible solution may be in the development of *intelligent* regionalization and/or extrapolation techniques, possibly based on hierarchical systems.

Lastly, to determine the reliability of landslide hazard and risk assessment models, errors associated with the thematic information used in the analysis has to be evaluated (Carrara et al., 1992). For the landslide dataset, the only currently available at the desired resolution for the entire Umbria-Marche territory, tests to check the reliability and degree of accuracy were carried out. For three sites, the Tescio and Carpina basins and the M. Coscerno area, the small-scale reconnaissance inventory was compared with detailed, large scale landslide mappings produced through the interpretation of aerial photographs of different vintages and scales, and extensive field surveys. Results are summarised in Table 5.

The reconnaissance inventory reports fewer landslides and approximately 50% less area affected by mass-movements. In the reconnaissance mapping the extent of the largest landslides was locally exaggerated and most of the smallest failures, of difficult interpretation on small scale aerial photographs within the time constrains of the project, were not identified. These results, associated with small scale geological and physiographic maps, may be used to correct the regional distribution of failures, allowing for more accurate regional landslide hazard and risk assessments.

CARPINA BASIN **67.22 kmq**	**Reconnaissance Inventory**	**Detailed Inventory**
Working Scale	25.000	10.000
Scale of Final Map	100.000	25.000
Scale of Photographs	33.000	13.000
Type of Photographs	Back & White	Colour
Total Landslide Area (kmq)	6.39	12.29
Percent of Basin	9.51	18.28
Number of Failures	108	1010
UNION (kmq)	14.77	
INTERSECTION (kmq)	3.91	
PERCENT ERROR	73.53	

TESCIO BASIN **59 kmq**	**Reconnaissance Inventory**	**Detailed Inventory**
Working Scale	25.000	10.000
Scale of Final Map	100.000	25.000
Scale of Photographs	33.000	13.000
Type of Photographs	Back & White	Colour
Total Landslide Area (kmq)	9.09	7.55
Percent of Basin	15.41	12.80
Number of Failures	-	-
UNION (kmq)	20.4	
INTERSECTION (kmq)	7.8	
PERCENT ERROR	61.76	

M. COSCERNO AREA **129.5 kmq**	**Reconnaissance Inventory**	**Detailed Inventory**
Working Scale	25.000	10.000
Scale of Final Map	100.000	25.000
Scale of Photographs	33.000	13.000
Type of Photographs	Back & White	Colour
Total Landslide Area (kmq)	9.95	18.22
Percent of Basin	7.68	14.07
Number of Failures	94	384
UNION (kmq)	20.44	
INTERSECTION (kmq)	7.77	
PERCENT ERROR	61.99	

Table 5. Errors associated with different landslide mappings in the Umbria-Marche region.

REFERENCES

Alexander, D. 1989. Urban landslides. Progress in Physical Geography. 13: 157-191.

Almagià R. 1910. Studi geografici sopra le frane d'Italia. Memorie Società Geografica Italiana, v. 13.

Antonini G., M. Cardinali, F. Guzzetti, P. Reichenbach, and A. Sorrentino. 1993. Carta Inventario dei Fenomeni Franosi della Regione Marche ed aree limitrofe.GNDCI publication 580 Map at 1:100,000 scale. In press.

Aronoff S. 1989. Geographical Information Systems: a management perspective. WDL Pubblications, Ottawa.

Brabb E.E. 1984. Innovative approaches to landslide hazard and risk mapping. Proceedings IV ISL, 1:307-323.

Brabb E.E. and B.L. Harrod (eds.). 1989. Landslides: extent and economic significance. Proceedings 28th International Geological Congress: Symposium on Landslides, Washington D.C., 17 July 1989, 385 p.

Burrough P.A., 1986. Principles of Geographical Information Systems for Land Resources Assessment. Claredon Press, Oxford, 194 p.

Carrara A. 1988. Drainage and devide networks derived from high-fidelity digital terrain models. in Chung C.F. et al. (eds) Quatitative analysis of mineral and energy resources, NATO-ASI Series, D. Reidel Pub. Co., Dordrecht, 581-597.

Carrara A., M. Cardinali, R. Detti, F. Guzzetti, V. Pasqui, and P. Reichenbach. 1991. GIS techniques and statistical models in evaluating landslide hazard. Earth Surface Processes and Landforms 20:427-445.

Carrara A., M. Cardinali, and F. Guzzetti. 1992. Uncertainty in assessing landslide hazard and risk. ITC Journal 1992(2):172-183.

DeGraff J. V. 1985. Using isopleth maps of landslide deposits as a tool in timber sale planning. Bulletin Association of Engineering Geology 22:445-453.

Guzzetti F., and M. Cardinali. 1990. Landslide inventory map of the Umbria region, Central Italy. 6th ICFL - ALPS 90, Milan, Italy, September 12, 1990. 273-284

Guzzetti F., and M. Cardinali. 1989. Carta Inventario dei Fenomeni Franosi della Regione dell'Umbria ed aree limitrofe. G.N.D.C.I. pub. n. 204. Map at 1:100,000 scale.

Guzzetti F., and M. Cardinali. 1991 Debris flows in the Central Apennines of Italy. Terra Nova 3:619-627.

Hansen, A. 1984. Landslide hazard analysis. Pages 523-602 in D. Brunsden and D.B. Prior (eds.). Slope Stability, John Wiley & Sons, New York.

Laurini R. and D. Thompson 1992. Fundamentals of spatial information systems. Academic Press, London, 680p.

R. K. Mark 1992. Map of Debris-Flow Prpbability, San Mateo County, California. U.S. Geological Survey, Miscellaneous Investigation Series, Map I-1257-M.

Schuster R.L., and R.W. Fleming. 1986. Economic losses and fatalities due to landslides. Bulletin Association of Engineering Geology 23:11-28.

Swanston D.N., and R.L. Schuster. 1989. Long-term landslide hazard mitigation programs: structure and experience from other countries. Bulletin Association of Engineering Geology 26:109-133.

Taylor F., and E.E. Brabb. 1986. Map showing landslides in California that have caused fatalities or at least $ 1,000,000 in damages from 1906 to 1984. U.S. Geological Survey. Miscellaneous Field Studies Map MF-1867.

UNESCO International Geotechnical Societies' Working Party on World Landslide Inventory. 1990. A suggested method for reporting a landslide. Bulletin International Association of Engineering Geology 41:5-12.

van Westen. 1993: Application of Gographic Information Systems to Landslide Hazard Zonation. ITC publication n. 15, 245 p.

Varnes D.J. and Commission on Landslides and Other Mass movement IAEG. 1984. Landslide hazard zonation: a review of principles and practice. Unesco Press, Paris. 63 p.