Proceedings of the 6th AGILE
April 24th-26th, 2003 – Lyon, France

TOWARDS THE CONSTRUCTION OF A SPATIAL DATABASE TO MANAGE LANDSLIDES WITH GIS IN MOUNTAINOUS ENVIRONMENT

Thiery Y.[1], Puissant A.[2], Beck E.[2], Malet J.-P.[1],

Remaitre A.[1],Maquaire O.[1], Sterlacchini S.[3]

[1] Institut de Physique du Globe, UMR 7516 - CNRS, 5, rue Descartes, 67084 Strasbourg Cedex, France.

[2] Image et Ville, UMR 7011 - CNRS, 3, rue de l'Argonne, Faculté de Géographie, 67000 Strasbourg, France.

[3] CNR-IDPA, Universita degli studi di Milano - Bicocca, Piazza della Scienza, 120126 Milano, Italy

1. INTRODUCTION

Geographic Information Systems (GIS) are more and more used for the management and the prevention of risks. These studies concern the preventive mapping of risk zones and the post-crisis management of catastrophes. The most recent progress concerns landslide hazard, as it has been the subject of many scientific researches for ten years, particularly in the development of susceptibility maps at medium (1/25,000) and large scales (1/10,000).

Three great methods based on GIS exist to analyze and find out the areas that gather the favorable conditions of predisposition to slope instabilities at these scales (Carrara *et al*, 1995 ; Soeters and Van Westen, 1996 ; Aleotti and Chowdhury, 1999):

- the qualitative approach, based on the knowledge of the cartographer (Kienholz, 1978; Leroi, 1996);
- the statistic approach (bivariate, multivariate), considered to be subjective (Carrara, 1988; Chung and Fabbri, 1993);
- the deterministic approach, taking into account geotechnical data to calculate parameters of slope stability (Wachal and Hudack, 2000).

The principal idea of these methods is that the pixel represents the "terrain unit". Thus, each movement is represented by "n" pixels. Each factor of predisposition to movements (lithology, topography, ...) is also represented by "n" pixels. The objective is to determine relations between these factors and landslides. When the relations are established by the way of different algorithms depending on the method, the results are validated for each pixel that belongs to a landslide or not. The final result represents a spatial correlation between the landslides distribution and each predisposition factor that is taken into account. This kind or approach permits to evaluate the probability that an event would occur in any region of the study area. It is consequently possible to point out the influence of the quality of the input data which can generate errors called "propagation errors" (Hunter and Goodchild, 1997 ; Aerts, 2002). More precisely, if the input data contain errors or if they are not very precise, it is difficult to evaluate the accuracy of the operations and of the final results (Aerts, 2002).

Each of these methods needs an input data set arranged in five categories (Soeters and Van Westen, 1996): geomorphology (landslides), topography (slope angle, slope aspect, elevation, …), geology (lithology, structure, superficial deposits), land-use (human activities, vegetation, infrastructures, …), hydrology (hydrographic network, swamps, …).

Studies have shown that the use of a similar data acquisition technique (by the analysis of aerial photographs) by different experts leads to distinct results. In this way, Van Westen *et al.* (1999) and Carrara *et al*. (1995) observed respectively 75% and 78% of difference in the localization of landslides, which was carried out by distinct persons on different areas. The expert knowledge of the scientist, the kind of data acquisition and their organization in a Database Management System (DBMS) influence the final results.

2. POSITION OF THE PROBLEM AND OBJECTIVES

While different GIS-based algorithms of hazard evaluation exist and are calibrated, few studies explain the way (1) to acquire the input data and (2) to structure the spatial database. For most of the studies, the first sources of information are aerial photographs and topographic maps. The first difficulty stands in the localization of instable zones on these documents (uneven topography, complex landslides, morphology hidden by vegetation). A field verification is therefore essential. The second difficulty stands in the integration of data that have different origins (image data, field observation, technical report, …) and unequal precision. Even though Soeters and Van Westen (1996) proposed lists of data necessary to be integrated, depending on the source documents and on the working scale, no evaluation of the data and any assessment of the methods of acquisition exist.
A methodological reflexion must be carried out on:

- the quality of the data, which will induce errors of localization and a loss of information (precision, accuracy), depending on the methods used to acquire the data;
- the structuring and the filing of the information in the spatial database, in order to avoid redundancies and to increase the interrogation speed.

The article shows that it is possible to decrease the subjectivity linked to the treatment of the input data by using a specific method developed to produce maps of susceptibility useful in the preventive management of the risk. The quality of the numerical simulation will depend on the way that data are acquired and stored (Flowerdew, 1991; Openshaw, 1991; Coppock, 1995; Burrough and Mc Donnell , 1998; Laarbi, 2000; Van Westen, 2000). In this context, a spatial database relating to factors of predisposition to landslides was built at a scale of 1/10,000. Data were then degraded (variation of resolution, generalization of classes, …) in order to define which optimal and reliable set of data would guaranty a statistical modeling (Carrara *et al.*, 1995) that would be close to reality. The susceptibility maps obtained with different sets of data were validated on a reference morphological map, produced by the expert method.

The optimal data set was then filed in a DBMS whose architecture was built according to three criteria: (1) opening of the system, (2) flexibility, and (3) easy management, what would improve the updating of data and the potential of the GIS to perform operations. The architecture was also defined according to three primary keys (location, landslides, activity), which were selected in order to make easier the search of data classified in tables (more than 10,000 referenced attributes). Different DBMS architectures were tested, especially the relational and object-oriented structures.

The study site is the Barcelonnette basin (surface: 90 km²), located in the Southern French Alps. This region is known for its slope instabilities and its complex morphology (Légier, 1977; Weber 2001).

3. INPUT DATA

The landslide susceptibility maps were built up by way of eleven factors of predisposition retained from the five categories of input data proposed by Soeters and Van Westen (1996) (Table 1). The database was set up in two stages. A first series of information layers were created only by the analysis of image data and topographic maps. The DTM was built up with the BDAlti ® coming from the IGN (resolution: 50m). Then, these layers were refined by integrating other sources of information progressively. A more precise DTM was derived from the digitizing of contour lines on topographic maps at a scale of 1/10,000. The data relative to landslides and to land-use were completed by field observations (precise boundaries of movements measured with a GPS, geological surveys, vegetation maps). Pieces of information learned from technical reports were also integrated into the spatial database. The single analysis of documents (maps and image data) did not seem to be sufficient for the construction of the database with precise and exhaustive data.

Categories of input data	Layers	Source of the information
Geomorphology	1. landslides	Aerial photographs, field survey, GPS, orthophotoplans, topographic map, morphological maps
Terrain geometry	2. slope 3. aspect 4. elevation	Topographic maps, BD Alti ®, new DTM
Geology	5. bedding 6. lithology 7. faults 8. superficial deposits	Geological map, field surveys, GPS, orthophotoplan, technical report
Land-use	9. land-use	Satellite images, field surveys
Hydrology	10. hydrographic network, 11. swamps	DTM, topographic maps Field surveys, topographic maps, GPS

Table 1 Main factors retained for the susceptibility assessment

4. RESULTS

4.1 Quality of input data

The data collected from different sources of information were compared to the field surveys and revealed many interpretation errors that concern essentially:

- the geomorphology, with problems of localization and recognition of several types of movements;
- the geology, with errors of nature and localization of some geological units and superficial deposits, problem with their thickness and superposition;
- land-use and particularly the forest cover, as the study area includes disturbed forests (very helpful for landslide recognition), which are not identifiable on aerial photographs or on satellite images.

The different tests realized on the different degraded data (rougher resolution, generalization of classes, use of better interpolation algorithms, …) revealed that the quality of the information as regards the boundaries of landslides, the topography, and the geology (lithology, superficial deposits) influenced the quality of the analysis and the final simulation. For the two first layers, the spatial precision of the data must therefore be as accurate as possible (precision inferior to 10 m, for a resolution of 10 m, slope angle classes every 10°).

For the superficial deposits, uncertainties remain concerning the classes of thickness: the retained classes were chosen depending on natural geological sections and were adapted to the topographic context. The retained classes of thickness are [0,5-2 m[; [2-5 m[; [5-10 m[, > 10 m.

For lithology, as it is hidden by the superficial deposits, the interpretation was performed from the geological map. Considering the uncertainties due to this interpretation, the spatial precision will be less accurate than for the landslides and the superficial deposits layers (precision of the boundaries tolerated until 20 m).

Table 2 presents the methods used for the data acquisition, the precision of the obtained data after errors and quality controls.

Layers	Methods of acquisition	Precision/acceptable error
1. landslides	1. Aerial photographs, , topographic map, morphological maps, geological map 2. Field survey, GPS, orthophotoplan	Less than 10 m
2. slope angle	1. BD Alti ®, 2. New DTM 3. Determining of classes with the GIS	Classes of slope angle every 10° minimum
3. aspect		8 classes of slope aspect
4. elevation		Pixel resolution: 10 m
5. bedding	1. Geological map 2. Field surveys	Acceptable error : between 5 and 15°
6. lithology	1. Geological map, morphological maps, aerial photographs, , topographic map, 2. Field survey, GPS, orthophotoplan	Between 10 and 20 m, considering the uncertainty due to the interpretation of the map
7. faults	Geological map	Less than 100 m
8. superficial deposits	1. Aerial photographs, , topographic map, morphological maps 2. Field survey, GPS, orthophotoplan	Less than 10 m for the boundary
9. land-use	1. Satellite images 2. Field surveys	Acceptable error up to 30 m (limited by the resolution of satellite images)
10. hydrographic network,	1. DTM 2. Topographic maps	
11. swamps	1. Topographic maps 2. Field surveys, GPS	10 m for the boundary

Table 2 Precision of data and methods of acquisition

4.2 Integration of the data in a GIS and implementation of DBMS

The integration of the optimal data set into a spatial database was done in three stages: (1) integration of the different sources of data, (2) integration of the attributes associated to each layer, (3) implementation of the database.

1.Integration of the different sources of data

The different layers were obtained by digitizing with ArcView ®, and were managed with ArcInfo ®.

For example, the acquisition of the data related to superficial deposits generated problems of superposition. The problem appeared during the rasterization: which information should contain the pixel? The information of the polygon of the A formation or the polygon B? ArcView ® does not manage this kind of superposition directly, two solutions are possible:

A first solution consists in digitizing each superficial deposit in different layers. This solution allows to take into account all the information relative to superficial deposits for the final calculation, but increases the time of acquisition and the number of data.

The second solution consists in integrating the different superficial deposits in ArcInfo ® after having digitized them in the same layer. As a matter of fact, ArcInfo ® manages the complexity of the superposition and creates a “common” polygon for the information of both superposed entities. The cartographer defines the order of the superposition (formation A over formation B or the opposite). This second solution provides a gain of time and avoid manipulating and managing too many different files.

2.Architecture of the database

The architecture of the DBMS was chosen considering three criteria: opening of the system, flexibility, easy management. The final choice concerns an architecture like relational, which has the advantage of being able: (1) to intersect files of different nature (polygons, polylines, points) by associating them in common tables by the way of common attributes, (2) to create new attributes like indexes from existing attributes, and to integrate them into the database, (3) to support modelling whereas other models like object-oriented ones do not support it so well (Burrough and Mc Donnell, 1998, Kunitoshi T. et Yutaca K., 2001).

Another advantage of this kind of DBMS is its new hybrid solution, which permits to enlarge the field of queries and new intersections of information (Healey, 1991; Burrough and Mc Donnell., 1998).

3.Choice of the attributes

The choices of the attribute data were done by taking into account literature and the study area characteristics. They were validated during the previous stage that determined the quality of the data. Each layer will contain in its relational database: an identity code (ID), common attributes like the closest place, its x and y coordinates (centroid for polygons, e for lines) and specific attributes (like indexes of form) derived from the source data.

5. CONCLUSION

The fulfilment of susceptibility maps with GIS at a scale of 1/10,000 in a complex environment requires to pay particular attention concerning (1) the type of data that would be introduced into the GIS and the database, (2) the way of acquiring these data, (3) the degree of precision needed for these data, and finally (4) how to structure the obtained information. The reasoning is not new, however, the means that are implemented here to decrease the subjectivity and the uncertainties linked to the particular context of the study area consist of a real progress in the methodology of the construction of a spatial database. Nevertheless, uncertainties and errors remain. It is absolutely necessary to be able to identify them in order to locate the origin of errors of further simulations.

6. ACKNOWLEDGEMENTS

The research carried out with the support of the European Research Program ALARM (Assessment of Landslide Risk and Mitigation in Mountains Area)

Mr. J. Hirsch and Mr A. Bernard are gratefully acknowledge for their help.

7. REFERENCES

[1] Augier C., Rebmann T (1992), Etude et cartographie géomorphodynamiques des risques naturels affectant la commune d'Enchastrayes – Ubaye – Alpes de Haute-Provence, Mémoire de Maîtrise, ULP, Strasbourg I, pp. 185.

[2] Aerts J. (2002), Spatial decision support for resource allocation, Integration of Optimization, Uncertainty analysis and Visualization techniques, PhD. Thesis, Univesiteit van Amsterdam, pp.153.

[3] Aleotti P., Chowdhury R. (1999), Landslide hazard assessment: summary review and new perspectives, Bulletin of Engineering Geology Environnment, 58, pp. 21–44.

[4] Burrough P.A., Mc Donnell R.A. (1998), Principles of Geographical Information Systems, Oxford University Press, pp. 333.

[5] Carrara A.(1988); Landslide hazard mapping by statistical methods. A black "box approach", The proceedings of the Workshop on Natural Disasters in European Mediterranean Countries, Italy, pp. 208-224.

[6] Carrara A., Cardinali M., Guzzetti F., Reichenbach P. (1995), GIS technology in mapping landslide hazard in Carrara A., Guzetti F (eds,), Geographical Information Systems in Assessing Natural Hazards, pp.135-176.

[7] Chung C.F., Fabbri A.G. (1993), The representation of geoscience information for data integration, Nonrenewable Resources, v.2:2, pp. 122-139.

[8] Coppock J.T. (1995), GIS and natural hazards: an overview from a GIS perspective, in Carrara A., Guzetti F (eds,), Geographical Information Systems in Assessing Natural Hazards, pp.21-34.

[9] Flowerdew R. (1991); Spatial data integration, in Maguire D.J. Goodchild M.F., Rhind D.W (eds), Geographical Information System, Vol. 1: Principles, pp. 375-387.

[10] Healey, R.G (1991), Database Management systems, in Maguire D.J. Goodchild M.F., Rhind D.W (eds), Geographical Information System, Vol. 1: Principles, pp. 251-267.

[11] Hunter G.J., Goodchild M.F. (1997), Modeling the uncertainty of slope and aspect estimates derived from spatial databases, Geographical Analysis, 29, pp. 35-49.

[12] Kienholz H. (1978); Maps of Geomorphology and natural hazard of Griend delward, Switzerland, scale 1/10 000; Artic and Alpine research, 10, pp. 169-184.

[13] Kunitoshi T. , Yutaka K. (2001), Considerations on Object-Oriented DBMS architecture, IPSJ SIG notes database system bastarct n° 094-014.

[14] Laarbi A. (2000), SIG et analyse multicritères, ed. Hermès, Paris, pp. 190.

[15] Leroi, E., 1996. Landslide hazard-risk maps at different scales: Objectives, tools and developments. In: Senneset, K. (ed.), Landslides, vol. I, Balkema, Rotterdam, pp. 35-51.

[16] Légier A. (1977), Mouvement de terrain et evolution récente du relief dans la region de Barcelonnette (Alpes de Hautes-Provence), PhD. Thesis, Grenoblel University, pp. 163.

[17] Openshaw S. (1991), Developing appropriate spatial analysis method for GIS, in Maguire D.J. Goodchild M.F., Rhind D.W (eds), Geographical Information System Vol. 1: Principles, pp. 389-402.

[18] Soeters R., Van Westen C.S. (1996), Slope instability, recognition, analysis, and zonation, in Turner A.K., Schuster R.L. (eds), Landslides investigation and mitigation, Transport Research board, National Research Council, special report 247, pp.129-177

[19] Van Westen C.J.,. Seijmonsbergen C., Mantovani F. (1999), Comparing landslide hazard maps, Natural Hazards 20, pp. 137–158.

[20] Wachal D. J. and Hudak P. F. (2000); Mapping landslide susceptibility in Travis County, Texas, USA, GeoJournal 51, pp. 245–253.

[21] Weber D. (2001), Contribution de la géomorphologie à la connaissance des mouvements de terrain dans les “Terres noires” alpines: le glissement-coulée de Super-Sauze (Alpes de Haute-Provence, France), PhD. Thesis, ULP, Strasbourg I, pp. 311 et annexes.