Test of Fuzzy Logic Rules for Landslide Susceptibility Assessment
Yannick Thiery, Jean-Philippe Malet, Olivier Maquaire

To cite this version:
Yannick Thiery, Jean-Philippe Malet, Olivier Maquaire. Test of Fuzzy Logic Rules for Landslide Susceptibility Assessment. SAGEO 2006: Information Géographique: observation et localisation, structuration et analyse, représentation, Sep 2006, Strasbourg, France. 16 p. hal-00276818

HAL Id: hal-00276818
https://hal.archives-ouvertes.fr/hal-00276818
Submitted on 2 May 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Test of Fuzzy Logic Rules for Landslide Susceptibility Assessment.

Yannick Thiery * — Jean-Philippe Malet * — Olivier Maquaire *

* CNRS UMR 6554, LETG – Geophen, Esplanade de la Paix, BP 5183, F-14032 Caen Cedex, France.
yannick.thiery@unicaen.fr; jeanphilippe.malet@unicaen.fr; olivier.maquaire@unicaen.fr;

ABSTRACT. Landslide Susceptibility Assessment (LSA) is defined as the spatial probability for a landslide to be generated in an area for many environmental factors. Currently, two approaches are used: (i) the qualitative approach based on expert opinion and knowledge of the relationship between the observed phenomenon and some predisposing factors and (ii) the statistical approach based on the statistical analysis of the relationship between the observed landslide and some predisposing factors. This paper proposes an exploratory attempt to use Fuzzy Logic Rules for mapping landslide susceptibility. The technique allows to describe the role of each predisposing factor (predictive variable) and their optimal combination. The best predictive variables identified by Fuzzy Logic are then introduced in a statistical bivariate model. The simulated maps obtained by both approaches are then compared and evaluated with an expert map, build with the prescribed rules of the French PPR (Plan de Prévention des Risques) methodology, and considered as a map of reference.

KEYWORDS: susceptibility, landslides, GIS, spatial analysis, fuzzy logic.
1. Introduction

Landslide susceptibility assessment (LSA) is defined as the spatial probability for a landslide to be generated in an area. It is deduced from the spatial correlation predisposing terrain factors (slope, landcover, superficial deposits, etc…) considered as predictive variables (PV) and the distribution of observed landslides within the territory considered as the response variable (RV). Therefore, LSA does not refer to the time dimension of the phenomenon (Sorriso Valvo, 2002). Several approaches may be used for LSA at coarse scale (from 1/50,000 to 1/10,000) like the qualitative approach based on expert knowledge and the statistic approach by bivariate or multivariate techniques. A description of these approaches with their advantages and limits can be found in Carrara et al. (1995), Soeters and van Westen (1996), Aleotti and Chowdury (1999) and van Westen et al. (2006).

Among the qualitative approaches, the Fuzzy Logic technique is based on subjective judgement about the relative importance of the predictive variables and their various states (Bonham-Carter, 1994). Some recent studies (Binhagi et al., 1998; Pistocchi et al., 2001; Ercanoglu and Gokceoglu, 2002; Tangestani, 2004) have demonstrated the flexibility of the approach for LSA either when the study area is large, when insufficient data is available, or when insufficient knowledge on the location and characteristics of the landslides is available. Moreover the technique allows to describe the role of each PV class, to identify the weight to assign to each PV class and to clarify the best combination of variables to introduce in the inference model.

This paper presents an exploratory attempt on the use of Fuzzy Logic rules for LSA in complex mountainous environments. The methodology is divided in three steps. First, the fuzzy memberships are defined from the knowledge between the observed landslides and each PV class; this step allows to describe the influence of each PV class on the location of landslides. The second step focuses on the “fuzzy inference network”, e.g. the role of the tested fuzzy operators on the results. The third step introduces the best combination of PV (identified by Fuzzy Logic) in a statistical bivariate model. Both maps are then compared to a qualitative map produced by expert knowledge (Plan de Prévention des Risques, PPR, methodology, MATE/METL, 1999). The research is applied on the North-facing hillslope of the Barcelonnette Basin (South French Alps) affected by numerous landslides and characterized by a complex morphology (Flageollet et al., 1999; Maquaire et al., 2003; Thiery et al., 2005; Malet et al., 2005).

2. Methodology

2.1. Background
In classical theory, the membership of a variable is defined on a binary scale (true = 1 or false = 0). In the fuzzy approach, the membership is expressed on a continuous scale from 0 (full non membership) to 1 (full membership). The membership function can be expressed as:

\[ \mu_A(x) : X \rightarrow [0, 1] \]  

where \( x \) is the variable, \( A \) is the fuzzy membership function and \( X \) is the universe of discourse in the interval \([0, 1]\) (Zadeh, 1965). The grade of membership reflects a kind of ordering that is not based on probability but on admitted possibility. Generally a low value (0 or near 0) is accorded for objects or classes which do not belong to the fuzzy set. Inversely, the grade of membership is large (1 or near 1) for objects or classes which fully belong to the fuzzy set (Carranza and Hale, 2001). Thus, based on expert opinion, individual classes of variables can be evaluated regarding their membership in a fuzzy set. The membership always reflects a general hypothesis. In this study, the general proposition for the predictive variables PV is “find favourable location of landslides”. For categorical variables, membership is not always linear but is related to a subjective hypothesis. For instance, for the categorical variable “lithology”, the fuzzy membership can be expressed with equation [2]:

\[ I = \{(x, \mu_l(x)) | x \in X\} \]  

where \( \mu_l(x) \) defines a grade of membership of lithology \( x \) in the class “favourable lithology for landslide location”.

2.2. Integration of Fuzzy Sets

Different operations allow to integrate the membership values (Bonham-Carter, 1994). Among them, the fuzzy algebraic operator PRODUCT, the fuzzy algebraic operator SUM and the fuzzy \( \gamma \)-operator combine the effects of two or more membership values in “blended” results (Carranza and Hale, 2001). Thus, each membership value has an effect on the output values (Bonham-Carter, 1994).

The fuzzy algebraic PRODUCT is defined as:

\[ \mu_{\text{combination}} = \prod_{i=1}^{n} \mu_i \]  

where \( \mu_i \) is the fuzzy membership function for the i-th map, and \( i = 1,2,\ldots, n \) maps are to be combined.

The fuzzy algebraic SUM is complementary to the fuzzy algebraic product, and is defined as:

\[ \mu_{\text{combination}} = 1 - \prod_{i=1}^{n} (1 - \mu_i) \]
where $\mu_i$ is the fuzzy membership function for the i-th map, and $i = 1, 2, \ldots, n$ maps are to be combined.

The fuzzy “$\gamma$-operator” is a combination of equation [3] and equation [4] (Zimmerman and Zysno, 1980), and is defined as:

$$P_{\text{combination}} = \left( \prod_{i=1}^{n} \mu_i \right)^{1-\gamma} \left( 1 - \prod_{i=1}^{n} (1 - \mu_i) \right)^{\gamma}$$  \[5\]

where $\mu_i$ is the fuzzy membership function for the i-th map, $i = 1, 2, \ldots, n$ maps are to be combined and $\gamma$ is a parameter between 0 and 1. When $\gamma$ is 1, the combination is the same as equation [4]. When $\gamma$ is 0, the combination is the same as [3]. A review of the role, advantages and limits of each operator is detailed in Bonham-Carter (1994).

2.3. Methodology

The methodology is split in three steps:

- First, a sensitivity analysis on the fuzzy membership values is performed in order to identify the predictive variables and their weights (reflecting some expert knowledge).

- Second, the integration of the fuzzy membership values in the inference model is evaluated by intermediate hypotheses. The intermediate hypothesis $H_1$ combines two or more PV in order to obtain a neo-predictive variable (NPV) representing a combination of predisposing factors favourable to landslide locations. This NPV can increase the predictive power of the maps (van Westen et al., 2003; Thiery et al. submitted). The intermediate hypothesis $H_2$ represents the combination between the NPV and the other PVs, and reinforces $H_1$ (Fig. 1). Several combinations of predictive variables are tested. The final membership values are then analysed with a cumulative curve representing the cumulative area versus the final membership values, and a susceptibility map is produced in four classes (S1, null susceptibility; S2, low susceptibility; S3, moderate susceptibility; S4, high susceptibility). The simulated maps are compared by calculating a relative error $\xi$, which relates the number of pixels of observed landslides to the number of pixels simulated in the S4 susceptibility class [6]:

$$\xi = \frac{o_v - p_v}{o_v}$$  \[6\]

where $\xi$ is the relative error, $o_v$ the observed value (e.g. number of pixels of the triggering zone of observed landslides) and $p_v$ the predicted value (e.g. number of pixels simulated in the S4 susceptibility class).
Third, the best dataset of PV identified by Fuzzy Logic is introduced in a statistical bivariate model based on Weight of Evidence (Bonham-Carter, 1994). The best Fuzzy Logic map and the statistical map are then compared and evaluated.
with a qualitative map produced by expert opinion (Malet et al., in press). Five statistical tests are calculated to compare the maps (e.g. correct classification rate, misclassification rate, sensitivity, specificity, kappa $K$ coefficient; Fielding and Bell, 1997).

![Figure 2](image-url)

**Figure 2.** *Test area located on the North-facing slope of the Barcelonnette Basin.*

### 3. Study area and description of the dataset (response and predictive variables)

#### 3.1. Study area

The study areas concerns the North-facing slope of the Barcelonnette Basin located in the French South Alps. The area extends over 11 km$^2$ (Fig. 2). The summit crests are capped by limestones and sandstones rocks; below the lower slopes are made of less resistant callovo-oxfordian black marls and are covered by moraine deposits (77% of the surface) and by forests (62%). The slope angles of these lower slopes are comprised between 5° and 45°. Various predisposing factors including lithology, tectonics, climate and the evolving land use have given rise to numerous slope movements on these slopes like rotational and translational landslides, mudslides and debris flows (Flageollet et al., 1999). The characteristics and the activity of these slope movements have been studied during the last ten years by several research teams (van Asch and Buma, 1997; Flageollet et al., 1999; Buma, 2000; Maquaire et al., 2003; Malet et al., 2005; Thiery et al., 2005).
3.2. Dataset of response and predictive variables

The variables used in this work can be grouped in one class of response variable (observed landslide, e.g. landslide inventory) and four classes of predictive variables (terrain geometry e.g. slope gradient, slope curvature, etc; terrain geology and geomorphology e.g. lithology, structure; superficial deposits, thickness of superficial deposits; terrain hydrography e.g. buffer of 100m around the main streams; and landcover.

The predictive variables were produced by crossing several source of information at 10x10m like topographic, geomorphological and geological maps, analysis of aerial photographs, analysis of satellite imageries and field survey (Thiery et al., 2004; Thiery et al., 2005 and Malet et al., in press).

Table 1. Characteristics of the main active landslide types observed in the Barcelonnette Basin, and main predictive variables (PV) according to expert knowledge. SLO: slope gradient, SF: superficial deposits; LIT: lithology; LAND: landcover; HYD: hydrology.

<table>
<thead>
<tr>
<th>Landslides type</th>
<th>Characteristics: location/geology</th>
<th>Main predictive variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Translational slide</td>
<td>Gentle slopes / At the contact moraine deposits/bedrock or in the bedrock</td>
<td>SLO, SF, LIT, LAND</td>
</tr>
<tr>
<td>Rotational slide</td>
<td>Along stream banks / Moraine deposits</td>
<td>HYD, SLO, SF, LAND</td>
</tr>
<tr>
<td>Shallow translational slide</td>
<td>Along stream banks / Moraine deposits and weathered black marls</td>
<td>HYD, SLO, SF, LAND</td>
</tr>
</tbody>
</table>

The inventory of landslides has been compiled at 1/10,000 scale. The standard procedure to inventory landslide events associates aerial photo-interpretation, field survey and analysis of local documents (Soeters and van Westen, 1996). Aerial photo-interpretation was carried out on photographs from IGN (Institut Géographique National) from year 2000 with at a proximal scale of 1/25,000. The photo-interpretation was completed by fieldwork to update information on the observed landslides. Three types were defined according to their morphological features (Dikau et al. 1996) and their predisposing factors. Table 1 shows the different landslides type and their characteristics. In this work, only the translational type of landslides is considered (Fig. 3).

4. Results

4.1. Generation of Fuzzy Sets
The fuzzy sets were generated on a test area representative (Fig. 2) of the complex geomorphology of the Barcelonnette Basin (Thiery et al., 2005; Thiery et al., submitted). Based on expert opinion and some statistical features (Fig. 3), different fuzzy memberships have been assigned for each PV classes. Figure 4 indicates the membership values. The different membership values are always comprised between 0.1 and 0.9 because it is not possible to be certain that the different PV classes are completely unfavourable or favourable for the occurrence of landslides.

Figure 3. Characteristics of the translational landslides observed in the the Barcelonnette Basin.

Two exceptions are made for the slope class [0°-5°] and for the landcover. For the slopes comprised between 0° and 5°, the membership value is 0 because no landslide has been observed on these slope gradients and because no landslide can occur on these slopes from a geomechanical viewpoint. For the landcover, Figure 3 indicates that majority of the translational landslides occur under forests; however, (i) forests can exhibit a stabilizing effect especially for shallow landslides (Greenway, 1987) and (ii) even if the other landcover classes are seldom represented they can influence landslide occurrence. Therefore, to identify the best membership value to assign to the landcover classes, some simulations were performed by assigning different membership values to the landcover class “forest”, “pasture”, “grassland”, “bare rock” and “black marl” and by keeping at a constant value the
other PV widely adopted by the scientific community for LSA (slope gradient, superficial formations and lithology). The membership value of the landcover classes “arable land”, “urban fabric” and “alluvial deposit” are defined by expert knowledge because no evidence relation can be hypothesized between them and landslide location (Fig. 5).

The integration of the fuzzy membership values is then performed with the fuzzy algebraic operator PRODUCT [3] which expresses the assumption that the combined membership values for the hypothesis must be present together for the hypothesis to be true (Carranza and Hale, 2001). The results are analyzed by calculating the relative error $\xi$ as described in Section 2.3. Figure 5 shows the variation of $\xi$ for different combinations of membership values for the landcover predictive variable. For instance, for the class “forest” a threshold is observed from Simulation 4 (corresponding to the membership value 0.4); therefore, this membership value is retained for the class “forest” for the next simulations.
4.2. Identification of the best inference model of Fuzzy Sets (combination of predictive variables)

Table 2 indicates some calculation results with different combinations of membership values and fuzzy operators. Among the different combinations, the inference model N°6 is considered as the best for LSA with a relative error $\xi = 0.26$. This inference model therefore associates:

- First, the fuzzy algebraic operator SUM between slope gradient and superficial deposits to verify the intermediate hypothesis H1 and to create the neo-variable NPV3.
- Second, the fuzzy $\gamma$-operator ($\gamma = 0.975$) between the variables NPV3, LIT and LAND to verify the intermediate hypothesis H2.

For H1, the operator SUM is more appropriate because two or more membership values complement one or two others (Carranza and Hale, 2001). Indeed, the class “moraine deposit” and different slopes gradient between 15° and 40° characterize the main geomorphological features of the triggering zones of translational landslides. This is manifested by four classes with high membership values either: (i) slope gradient between 20° and 35° with moraine deposits (membership value of 0.99), (ii) slope gradient between 35° and 40° and moraine deposits (membership value of 0.98); (iii) slope gradient between 35° and 45° and weathered black marls (membership value of 0.96) and (iv) slope gradient between 15° and 20° and moraine deposits (membership value of 0.95).

Figure 5. Trial & Error test on the identification of the membership value to assign to the predictive variable “Landcover”. Computations are performed with SLO, SF and LIT variables kept at constant membership values (see figure 2 for the membership values).
Finally, despite the same relative error $\xi$ between the inference models N°5 and N°6, the N°6 is considered as better because the N°5 overestimates the high susceptible class (S4) with 55% of the total area.

The final susceptibility map produced by Fuzzy Logic Rules therefore simulates 25% of the area with high susceptibility (S4), 17% of the area with moderate susceptibility (S3), 3% of the area with low susceptibility (S2) and 55% of the area with null susceptibility (S1).

<table>
<thead>
<tr>
<th>Simulation</th>
<th>Inference model and predictive variables used</th>
<th>Fuzzy operator used</th>
<th>Relative error $\xi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>N°1</td>
<td>NPV 4 + LIT + LAND</td>
<td>SUM</td>
<td>0.57</td>
</tr>
<tr>
<td>N°2</td>
<td>NPV 4 + LIT + LAND</td>
<td>PRODUCT</td>
<td>0.57</td>
</tr>
<tr>
<td>N°3</td>
<td>NPV 4 + LIT + LAND</td>
<td>$\gamma$-operator ($\gamma = 0.975$)</td>
<td>0.55</td>
</tr>
<tr>
<td>N°4</td>
<td>NPV 3 + LIT + LAND</td>
<td>PRODUCT</td>
<td>0.34</td>
</tr>
<tr>
<td>N°5</td>
<td>NPV 3 + LIT + LAND</td>
<td>SUM</td>
<td>0.26</td>
</tr>
<tr>
<td>N°6</td>
<td>NPV 3 + LIT + LAND</td>
<td>$\gamma$-operator ($\gamma = 0.975$)</td>
<td>0.26</td>
</tr>
</tbody>
</table>

Table 2. Relative error associated to the inference models tested. NPV 3 is obtained between SLO and SF by the fuzzy algebraic operator SUM. NPV 4 is obtained between SLO and SF by the fuzzy algebraic operator PRODUCT.

5. Discussion: comparison of the Fuzzy Logic susceptibility map with a statistical susceptibility map and a qualitative expert susceptibility map.

5.1. Application of a bivariate statistical model on the same dataset

The best dataset identified by Fuzzy Logic Rules (simulation N°6, Table 2) is introduced in a statistical bivariate model (Weight of Evidence, WofE). WofE is considered as one of the best procedure for LSA if the conditional problem between the predictive variables is tackled (Soeters and van Westen; 1996; van Westen et al., 2003). A description of the method and its applications of the Barcelonnette Basin are detailed in Thiery et al. (submitted). The statistical susceptibility map shows some similarities with the map simulated with Fuzzy Logic (Fig. 6). Nevertheless, as the WoFe model computes landslide occurrences according to the observed number of pixels of landslides, some locations are not recognized in the high susceptibility class (S4) by the statistical model (e.g. pastures and bare rocks).
Figure 6. Susceptibility maps of translational slides obtained by qualitative approach (A), Fuzzy Logic Rules (combination NPV 3, LIT; LAND; γ-operator 0.975) (B) and statistical Weight of Evidence model (C). (D) represents the aerial photograph from IGN (2000). (E) represents the location of the A, B, C and D.
5.2. Comparison with the qualitative map

Finally, the susceptibility maps produced by Fuzzy Logic Rules and by the WoFe bivariate model are compared to a qualitative expert map considered as the reference map (Table 3). The qualitative map was elaborated according to the French official methodology of PPR (Plan de Prévention des Risques; MATE/MATL, 1999). This methodology requires a large overview of the area to identify sectors with homogeneous environmental characteristics, and takes into account the possibilities of landslide developments for the next one hundred years.

The maps are compared with a confusion matrix (Fielding and Bell, 1997). The correct classification rates indicate good classification of the four susceptibility classes (between 0.78 and 0.80; Table 3). The sensitivity index is more informative as it represents the probability that a class S of the models is correctly classified in the same class of the qualitative map. The sensitivity coefficient is good for class S1 and class S4 but not for class S2 and class S3 (Table 3). This classification error is confirmed with the kappa $K$ coefficient. A value above $K = 0.4$ indicates a good fit between the maps. For both model, the $kappa K$ coefficients are good for class S1 and class S4 (Table 5), but the values are very low for class S2 and class S3.

<table>
<thead>
<tr>
<th>Fuzzy Logic Susceptibility map</th>
<th>Statistical Susceptibility map</th>
</tr>
</thead>
<tbody>
<tr>
<td>Susceptibility class General</td>
<td>Susceptibility class General</td>
</tr>
<tr>
<td>S1 S2 S3 S4</td>
<td>S1 S2 S3 S4</td>
</tr>
<tr>
<td>CCR</td>
<td>0,80 0,85 0,82 0,80 0,82 0,78 0,87 0,84 0,80 0,82</td>
</tr>
<tr>
<td>MCR</td>
<td>0,20 0,15 0,18 0,20 0,18 0,22 0,13 0,16 0,20 0,18</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0,81 0,19 0,13 0,62 0,64 0,77 0,13 0,13 0,63 0,64</td>
</tr>
<tr>
<td>Specificity</td>
<td>0,22 0,91 0,91 0,86 0,88 0,80 0,91 0,92 0,84 0,88</td>
</tr>
<tr>
<td>Kappa $K$</td>
<td>0,58 0,08 0,05 0,47 0,52 0,53 0,03 0,05 0,43 0,52</td>
</tr>
</tbody>
</table>

Table 3. Accuracy tests between the qualitative expert susceptibility map, the Fuzzy Logic susceptibility map and the statistical Weight of Evidence susceptibility map. (CCR: correct classification rate; MCR: misclassification rate).

For the Fuzzy Logic susceptibility map, these results may be attributed to the general hypothesis of the model “find favourable location of landslides”. In this case the membership values and their combinations have been assigned in order to find the most relevant susceptible areas. Thus, the Fuzzy Logic Rules tends to simulate a “binary” susceptibility map with high and null susceptibility classes.

For the statistical bivariate model, the same explanations can be hypothesized since the statistical algorithm is based on binary evidences, and has been elaborated
to recognize areas with identical environmental characteristics prone to landslides. Thus, the moderate (S2) and low (S3) susceptibility classes that may be interpreted by an expert through the association of the predisposing factors and taking into account the principle of precaution can not be estimated with Fuzzy Logic Rules or statistical models.

6. Conclusion

The application of Fuzzy Logic Rules to Landslide Susceptibility Assessment in a complex mountainous environment at 1/10,000 scale has shown its efficiency to recognise favourable landslide prone areas from non favourable landslide prone areas. The Fuzzy Logic model requires detail knowledge of landslide characteristics and of their relation with the predisposing factors. Indeed, the knowledge of the expert is of paramount importance to apply such type of model.

Proposing a two-steps inference model with intermediate hypotheses seems appropriate, and allows to combine the membership values with different fuzzy algebraic operators. In this study, the first intermediate hypothesis H1 which corresponds to the main geomorphological features influencing landslide location is obtained by the SUM operator. The second intermediate hypothesis H2 reinforces these features by local factors and the use of the fuzzy $\gamma$-operator. This combination increases the predictive power of the maps.

The best combination of predictive variables identified with Fuzzy Logic can then be introduced in a statistical bivariate model. This strategy allows to be free of long statistical procedures to decrease the conditional dependence problem inherent to the statistical approach (Thiery et al., submitted). Thus, the choice of the predictive variable may be easier for the expert in charge of the mapping.

However, the methods used in this study are not able to recognize the moderate (S2) and the low (S3) susceptibility classes because of the general hypothesis “find favourable location of landslides” used in our methodology. Therefore, only binary susceptibility maps (high susceptibility, null susceptibility) can be produced.

Acknowledgements

This research was supported by the European Union through the research programme ALARM (Assessment of Landslide Risk and Mitigation in Mountain Areas), Contract EVG1-2001-00018, 2002-2004, Coordinator: S. Silvano (CNR-IRPI, Padova).
7. Bibliography


Test of Fuzzy Logic Rules for Landslide Susceptibility Assessment


