Landslide susceptibility assessment in the hilly valleys of the Pays d'Auge using logistic regression (Normandie, France)

Mathieu Fressard, Yannick Thiery, Olivier Maquaire

First attempt of landslide susceptibility assessment in the hilly valleys of the Pays d'Auge plateau (Normandy, France)

M. Fressard & O. Maquaire
UMR 6554 LETG-Géophen, University of Caen Basse-Normandie, France

Y. Thiery
GSC consultant, Paris, France
UMR 6554 LETG-Géophen, University of Caen Basse-Normandie, France

ABSTRACT: In Europe, many studies about landslides have been performed in mountainous environments. However, a large proportion of sloping hilly valleys in Western Europe are also affected by slope instabilities. This paper presents a first attempt of landslide susceptibility mapping on a selected representative area of 24 km² located in the Pays d'Auge plateau in Normandy (France). The main objective is to define a quick reproductive indirect mapping technique at 1:10.000 scale with a set of rapid available data. The technique could be used as an operational mapping technique. In this case, only shallow landslide susceptibility was assessed by the logistic regression technique. The conclusions show interesting results in terms of high susceptibility areas locations, nevertheless, the model performances can still be improved by the introduction of new dataset.

1 INTRODUCTION

Since the last two decades, landslides susceptibility assessment studies have mainly been performed in mountainous and coastal environments. However some hilly valleys of the North-West of Europe present a wide variety of landslides phenomena which constitute a proper risk for the exposed population (van den Eeckhaut et al., 2006; 2009). The Pays d'Auge plateau (Normandy, France) is one of those places largely affected by slope instability processes but still few have studied it (Fressard, 2009). These phenomena induce many damages to infrastructures like roads, buildings and pipes (Fig. 1).

Assessing landslide susceptibility with a minimum of set of data, a reproducible methodology and GIS techniques, is a challenge for earth-scientists, government authorities and resource managers (Glade and Crozier, 2005). In the Pays d'Auge plateau, the actual operational resources for landslide hazard mapping are insufficient given the landuse planning needs (lack of accuracy and incompatibility with the national risk prevention plans: PPR), (MATE/MATTL, 1999; Fressard, 2009). The first step of landslide hazard analysis is also to create landslide susceptibility maps. In order to be operational, these maps must be adapted to the scale of the existing national risk prevention plans i.e., 1:10.000 scale (MATE/MATTL, 1999). Indirect approaches, based on statistical conditional analysis, comparison of landslides inventories and predisposing factors (Carrara et al., 1995), appear as a good local alternative to the '*landslide susceptibility map'* proposed by the DIREN (*Regional Supervision of the Environment*, 2008) which is known for its imprecision (Fressard, 2009). The statistical approach has been the object of several studies and is often considered as the most objective method avoiding the problem of the expert subjectivity (Soeters and van Westen, 1996). Moreover, statistical methods are useful in landslide mapping, because a limited data set can produce meaningful results (van Westen, 2000).

This paper presents a first attempt of a susceptibility mapping at the 1:10.000 scale using logistic regression for the 24 km² test site of the '*Mont Saint Léger*' in the Pays d'Auge plateau.

Figure 1. Example of a damage induced by a landslide triggering in the Pays d'Auge plateau (Source: TP-Geo)

2 STUDY AREA AND LANDSLIDES

2.1 *General presentation and geomorphology*

The Pays d'Auge plateau is an agricultural region of Normandy of 2000 km² widely affected by instability processes like brutal subsidence and landslides. The hillsolpes are mainly covered by grasslands and forests. This region has a maritime temperate climate with a rainfall of ± 800 mm.yr^{-1} regularly distributed over four seasons.

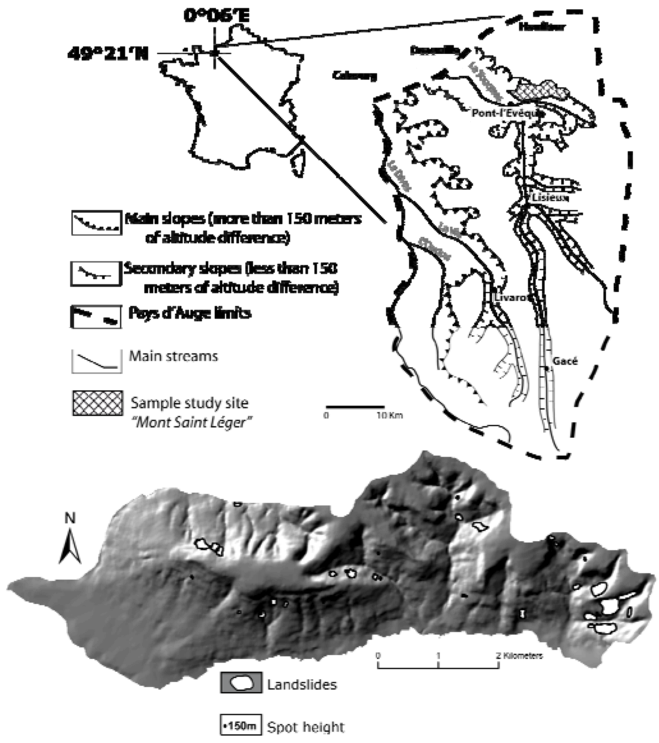


Figure 2. Study site location, landslides location and sample study area

The regional topography, lithology and hydrology are important environmental factors controlling slope stability (Helluin, 1988; Lautridou, 1971). Hillslopes are generally not very steep. Only 10% of the hillsolpe sections have slope gradients over 20% and 67% are ranked between 5% and 15%. In the late Tertiary and early Quaternary, differential erosion of lithological layers shaped the actual topography of the area. The lithology consists of a monoclinal structure based on two major entities: (1) an alternation of marly to marly-calcareous Jurassic layers and (2) a thick layer of aquifer Cretaceous chalks. The contact zone of these two major entities is composed of clayey glauconitic sands (Doré et al., 1977). The Pays d'Auge hillslopes are overlain by various types of allochthonous and autochthonous Quaternary deposits. The spatial distribution of these surficial deposits is relatively unknown (Fressard, 2009). The most represented formations are flint clays, altered marls, clays, sands and aeolian loess. Their thickness can vary between 2 to 5 meters (Elhaï, 1983).

2.2 *Landslides data*

The BDMvT database (http://www.bdmvt.net/) has been created in 1994 and is developed by the '*French Geological Survey'* (BRGM) in collaboration with the '*Ministry of Ecology, Energy, Sustainable Development and Spatial Planning'* (MEDAD), the '*Ministry of Education and Research*' (MESR), the '*Laboratoire des Ponts et Chaussées*' (LRPC) and the '*Service de Restauration des Terrains en Montagne*' (RTM). It is available for the department of Calvados since 2005. The database draws up an inventory of the past and active known landslides. For each event the landslide type, the location, the date, the activity, and sometimes the lithology affected are mentioned. Nevertheless, the BDMvT database is not exhaustive, and some uncertainties in terms of landslide type, date of occurrence and location are observed (Malet et al., 2009; Fressard, 2009). The occurrence of landslides events have been mapped at the municipality scale, i.e., 1/100.000 scale.

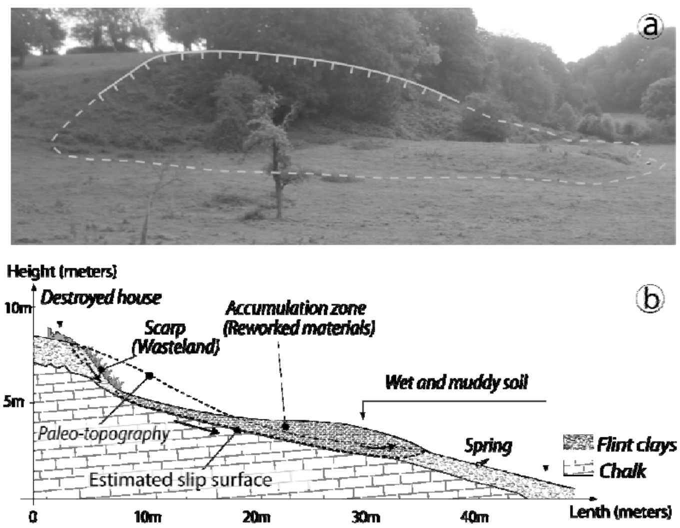


Figure 3. Example of a shallow translational landslide (a) and its interpretative slope profile (b)

The incompatibility between this database and the 1:10.000 scale is obvious. A new landslide inventory was performed for a selected representative 24 km² study site (Fig. 2). The landslide inventory was compiled at the 1:10.000 scale through air-photo interpretation with systematic field survey. The landslides were mapped in the field by cartographic GPS with a 5 meters precision. The landslides boundaries were classified into two zones: (1) landslide initiation zone and (2) landslide accumulation zone (Atkinson & Massari, 1998; van den Eeckhaut, 2006; Thiery et al., 2007). Morphological parameters, landslide type and state of activity were stored in a GIS database. Four major landslide types have been defined, according to the Cruden and Varnes (1996) classification: (1) shallow and very shallow translational landslides (Fig. 3), (2) shallow and very shallow rotational landslides, (3) rotational deep seated landslides and (4) solifluction processes.

A large part of the rotational deep seated landslides have been identified as naturally stabilized or inactive mature (Dikau et al., 1996). For the susceptibility analysis, only the active shallow landslides (Maquaire & Malet, 2006) were introduced in the analysis.

2.3 Landslides predisposing factors

One main objective of the study was to work with direct available data. Nevertheless, the information proposed by the National Height Elavation Database (BDalti®, *French Geographical Institute*) and the CORINE Land Cover database were not accurate enough to be used with the indirect approach at the 1:10.000 scale model (Thiery et al., 2004). Therefore, a new DTM (10-meters resolution and ± 2 meters precision) was created by the IDW interpolation technique on digitized contour lines of the 1:25.000 scale topographic map. The slope gradient, planar curvature an aspect maps were derived from the new DTM. A landuse map was obtained by air-photo interpretation on the 2006 orthophotography. Geological parameters were obtained by digitization of the 1:50.000 scale georeferenced national geological maps (Debrand-Passard et al., 1987; Guyader et al., 1970). During the digitizing, shapes corrections have been performed in order to increase the map accuracy.

3 METHODOLOGY

The logistic regression technique has been chosen for its simplicity and its robustness. Furthermore, the method is not sensitive to the conditional dependence of the introduced variables and has given good results in similar hilly environments (van den Eeckhaut et al., 2006; 2009).

3.1 Logistic regression principle

The logistic regression describes the relationship between a dichotomous response variable (*Y*, i.e., the presence or absence of landslides) and a set of predictive variables ($x_1, x_2, \ldots x_n$). The predictive variables may be continuous or discrete and do not need a normal frequency distribution. The logistic response function can be written as (Allison, 2001):

$$P(Y=1)=\hat{p}=\frac{1}{1+e^{-(\hat{\alpha}+\hat{\beta}_1 x_1+\hat{\beta}_2 x_2+\cdots+\hat{\beta}_n x_n)}}$$

Where $\hat{p}$ is the spatial probability of occurrence of a landslide, $\hat{\alpha}$ is the intercept $\hat{\beta}_i$ is the coefficient for the independent variables $\hat{x}_i$ estimated by maximum likelihood.

3.2 Modelling strategy

For this study, a 24 Km² sample area was selected. This site was judged representative of the majority of the Pays d’Auge hillslopes with a general South-West orientation and a woody to grassland landuse (Fressard, 2009).

The statistical model was implemented in ArcGIS 9.3.1® through the ArcSDM extension (Kemp at al., 2001).The proposed methodology consists in four major steps:

(I) The first step consists in the choice of the best predictive variables. Therefore, statistical tests were realized. Chi² test of association was performed. This test was completed by a *V* Cramer’s test because the Chi² is very sensitive to the number of introduced variables (Davis, 2002; Pistocchi, 2002; Thiery, et al., 2007). These tests of association aim to find a relationship between each predictive variable and the occurrence of landslides. Only the predictive variables and classes of predictive variables showing the highest coefficients of associations were selected by the expert.

(II) In order to preserve a set of data for the validation step, only 50% of the triggering zones cells were used for the model calibration. The other 50% were used for the validation step (Thiery et al., 2007). Successive model iterations were realized with a stepwise introduction of the predictive variables (Thiery et al, 2007).

(III) Each map was classified into four susceptibility classes as suggests the MATE and MATL 1999 i.e., null, low, moderate and high. Relative error ξ calculation (see formula on table 2) was performed between the highest susceptibility classes and the calibration and validation response variables. This test was completed by the prediction rate analysis proposed by Chung and Fabri (2003). This curve shows the percentage of the study area, ranked from most to the least susceptible, against the cumulative percentage of the area of the triggered landslides in each susceptibility class (Jaiswal et al., 2010). In order to obtain a robust validation of the models and as it is mentioned by Chung and Fabbri (2003), the success rate is performed with the 50% of the calibration triggering zone cells and the prediction rate is performed with the 50% of the validation triggering zone cells.The map showing the weakest relative error ξ and the best rate of predicted cells was selected as the best susceptibility map.

(IV) Finally, the reclassified maps were submitted to the expert opinion, which constitutes the last validation step.

4 RESULTS

As it is known by expert opinion, the geological structure has a very weak influence on shallow landslides processes (Fressard, 2009). In the Pays d'Auge valleys, shallow landslides are only affected by the surficial formations which are for the most part, independent of the geological structure. This variable was also removed from the analysis. Only the highest slope classes (5-10%; 10-15%; 15-20% and >20%) were added to the model according to the expert opinion confirmed by the statistical analysis. The Chi² and *V* Cramer's tests have permitted to identify only three influent landuse classes: grassland, forest and shrub land. The other variables (planar curvature and aspect gradient) were not reclassified to be introduced in the model.

INTRODUCED VARIABLES	CAL. MV ξ*	VAL. MV ξ**	EXPERT OP.***
(1) Slope-landuse	0.19	0.15	Not likely
(2) Slope-landuse-aspect	0.31	0.28	Not likely
(3) Slope-landuse-Planar curvature	0.27	0.21	likely
(4) Slope-landuse-Planar curvature-aspect	0.29	0.26	Not likely

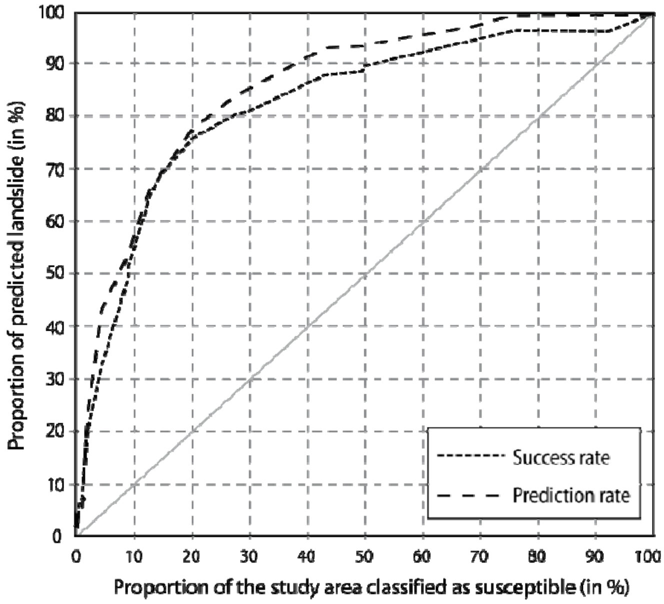
Relative error ξ = (total number of triggering zones cells-total number of high susceptibility triggering zones cells) / total number of triggering zones cells
*Relative error of the calibration variables
**Relative error of the validation variables
***Expert opinion

Table 1. Relative error and expert opinion of the different model iterations

The table 2 shows the results obtained for the different simulations realized with the stepwise insertion of the predictive variables in the model. The weakest relative error is obtained with only the slope and landuse variables. However the expert opinion has shown the lack of plausibility of this map. High susceptibility classes are over-estimated whereas null and low classes are in minority position. It was therefore not retained as the best map. According to the different simulations slope aspect was not retained as a significant predictive variable, the final susceptibility map was obtained with slope, landuse and planar curvature. The final susceptibility map (Fig. 5) shows good agreement with the landslide inventory map and is characterized by ξ values of 0.27 for the calibration variables and 0.21 for the validation variables. Finally, 20% of the study area is classified as high susceptibility, 30% as moderate and 26% as low.

The figure 4 shows the prediction and success rate curves for the best landslide susceptibility map. The shape of the two curves is very similar, which confirms that 50% of triggering zone cells are sufficient to obtain good results. For the two curves 20 % of the susceptible area are sufficient to predict ± 80% of the triggering zone cells.

Figure 4. Success and prediction rate analysis curve for



simulation 3 (final susceptibility map).

5 CONCLUSION AND DISCUSSION

This first attempt shows the usefulness of the logistic regression technique in terms of indirect susceptibility assessment at the 1:10.000 scale. This method allows calibrating rapidly a model which fits well with the landslide inventory without overestimating high susceptibility class. Nevertheless, the problem of the spatial resolution of the direct available data (BDMvt, BDAlti®, CORINE Land Cover) still remains (Thiery, 2004). New datasets that fits to the national risk prevention plan, i.e., 1/10.000 scale must be created and increase the time needed to perform the analysis.

The produced map has permitted to indentify high susceptibility zones on the "Mont Saint Léger" study site that strongly differs from the *'landslide susceptibility map'* proposed by the DIREN (*Regional Supervision of the Environment*, 2008). The shapes of the susceptibility classes are better defined and do not only give a high susceptibility score to the steepest parts of the slopes.

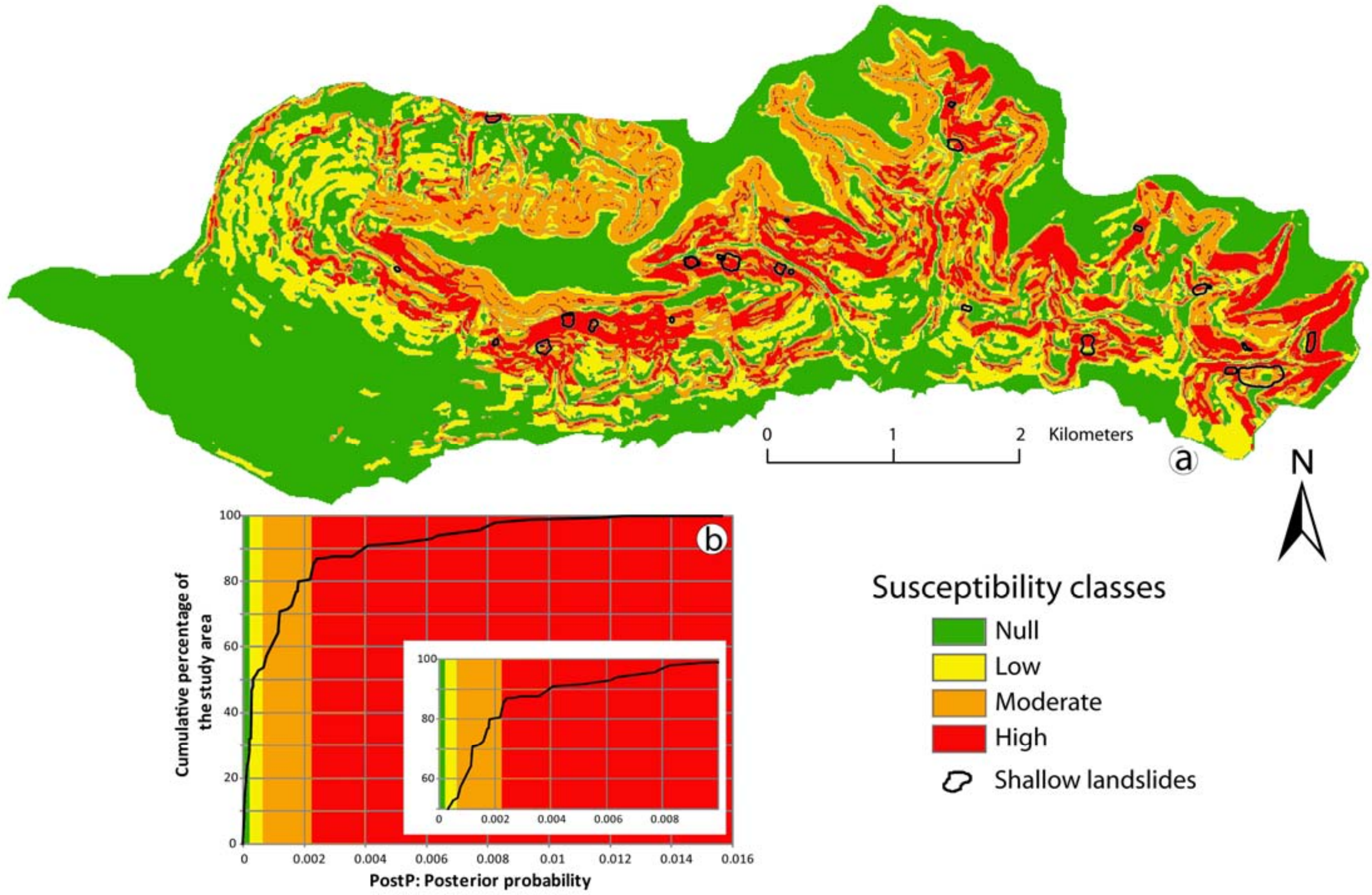


Figure 5. Final susceptibility map (a) and classification curve of the posterior probabilities (b)

The field campaigns had permitted to observe, that the majority of shallow landslides are initiated on less steep slopes, i.e. between. According to the model, the forest has a stabilization effect on the landslide susceptibility and a large part of the landslides are present on shrub land areas. This map could also be also a good alternative to the actual operational documents to map the shallow landslide susceptibility in this hilly context.

However, this method only provides a spatial probability and ignores the temporal aspect of the landslide triggering (van Westen et al., 2006). The probabilities obtained are used as a score to identify susceptibility zones and should not be considered as a true occurrence probability.

Even if the model has shown good results on the test study site, it still remains difficult to transpose the results to a larger site. Different reasons can explain these transposition problems:

(I) From an expert point of view, some very important data are still missing. Surficial formations are known to be very important in the shallow landsliding process (Maquaire & Malet, 2006). Due to the complexity of their distribution in the study area, their mapping is known to be a difficult and long process, and available documents still do not exist on this site (Elhaï, 1983; Fressard, 2009). It was also chosen not to use this variable in a first step.

(II) Moreover the presence of perched ground-water shows the interest of this particular geological configuration which generates a permanent water discharge on mild slope. The surficial formations and soils are also permanently wet, which could be an important landslide predisposition factor. The creation of a soil wetness index should be an interesting input in the landslide susceptibility assessment.

The introduction of these two predictive variables appears essential in the model improvement process, and shows that landslide susceptibility assessment can hardly been performed at the 1:10.000 scale without a robust geomorphologic analysis (Thiery et al., 2007; van Westen et al., 2003; 2008). However, the time needed to perform this complex approach makes it difficult to be applied as an operational method.

(III) In cases of rare events, logistic regression can over or underestimates the final spatial probabilities. To avoid these problems, correction calculations can be integrated to the ordinary logistic regression process (van den Eeckhaut, 2006). This method will soon be tested in the Pays d'Auge plateau.

These firsts results of indirect susceptibility mapping show good results on the "*Mont Saint Léger*" test study site. This final map clearly better identifies landslides than the current available operational predisposition map. Logistic regression can also be an interesting alternative of quick indirect operational mapping, even if the problem of the input data is still present. These results can clearly be improved by the introduction of complementary predictive variables and statistical correction tests.

6 REFERENCES

Allison, P.D., 2001. *Logistic Regression Using SAS system: Theory and application*. Willey Interscience, New York. 288pp.

Atkinson, P.M., et Massari, R., 1998. Generalized linear modelling of landslide susceptibility in the Central Apennines, Italy. *Computers & Geosciences*, v. 24, no. 4, p. 373-385.

Carrara, A., Cardinali, M., Guzzetti, F., Reichenbach, P., 1995. GIS technology in mapping landslide hazard. *Geographical information systems in assessing natural hazards*,, p. 135-175.

Chung, C.J.F., et Fabbri, A.G., 2003, Validation of spatial prediction models for landslide hazard mapping: Natural Hazards, v. 30, no. 3, p. 451-472.

Couëffé, R., Arnaud, L., Choutier, J.P., Lebert, J.F., Pasquet, V., Hugot, V., 2005. Inventaire préliminaire des glissements de terrain du Calvados (Basse-Normandie). Rapport final. *Rapports publics du BRGM*. 149p.

Cruden, D.M., et Varnes, D.J., 1996. Landslide types and processes. *Landslides, Investigation and Mitigation*,, p. 36–75.

Davis, J.C., 2002. *Statistics and Data Analysis in Geology*, Third Edition. John Wiley & Sons, New York, 638 p.

Debrand-Passard, S. Prost, A.E., Goyallon, J. 1987.Carte géologique de la france au 1 :50.000ème et livret explicatif, feuille de Lisieux. *Editions du BRGM*, France.

DIREN, 2008. Carte de prédisposition aux mouvements de terrain (glissements de pente, coulées de boue et fluage). *Les documents de la diren*. 5p.

Dore, F., Juignet, P., Larsonneur, C., Pareyn, M. Rioult, M., 1977. *La Normandie. Guides géologiques régionaux*. Paris, Masson. 360pp.

Elhaï, H., 1983. La Normandie occidentale : entre la seine et le golfe Normand-Breton. Etude morphologique. *Thèse d'état. Université Paris, Sorbonne*. Bière Bordeaux. 581p.

Fressard, M., 2009. Morphodynamique des versants du Pays d'Auge continental, Fonctionnement, heritages et risqué associés. *Memoire de master II recherche, Université de Caen Basse-Normandie*, 135pp.

Glade, T., et Crozier, M.J., 2005. A review of scale dependency in landslide hazard and risk analysis. *Landslide hazard and risk*, p. 75-138.

Guyader, J., Pareyn, C.L., Viallafond, L., Juignet, P., 1970. Carte géologique de la france au 1:50.000ème, feuille du Havre. *Editions du BRGM*, France.

Guzzetti, F., Reichenbach, P., Ardizzone, F., Cardinali, M., Galli, M., 2006. Estimating the quality of landslide susceptibility models. *Geomorphology*, v. 81, no. 1-2, p. 166-184.

Helluin, E., 1988. Les mouvements de terrain de février 1988 dans le Calvados. In Risques naturels et analyses pour une prévention. L'hiver 1987-1988 en Basse-Normadie. *Travaux du CREGEPE*. pp 8-10

Jaiswal, P., van Westen, C., Jetten, V., 2010. Quantitative landslide hazard assessment along a transportation corridor in southern India. *Engineering geology*, accepted manuscript.

Kemp, L.D., Bonham-Carter, G.F., Raines, G.L.,n Looney, C.G., 2001. Arc-SDM: Arcview extension for Spatial Data Modelling using Weights of Evidence, Logistic Regression, Fuzzy Logic and Neutral Network Analysis. http://ntserv.gis.nrcan.gc.ca/sdm/

Lautridou, A., 1971. Formations superficielles et dynamique des versants du Pays d'auge. *Excursions dans le Pays d'Auge, Journée du 9 juillet 1971*. 28p.

Malet, J.P., Thiery,Y., Hervàs, J., Günther, A., Puissant, A., Grandjean, G., 2009. Landslide susceptibility mapping at 1:1M scale over France:exploratory results with heuristic model. *In proc. of Landslide process, from geomorphologic mapping to dynamic modelling. A tribute to Dr. Theo van Asch*. Strasbourg france, feb. 2009. pp. 315-320

Maquaire, O. & Malet, J.P., 2006. Shallow landslideing. In Boardman, J. Poesen, J., (eds), *Soil Erosion in Europe*, Willy, Chapter 2.8, 583-598.

MATE/MATL, 1999. Plan de Prévention des Risques (PPR) : Risques mouvement de terrain. *Ministère de l'aménagement du territoire (MATE), ministère de l'équipement des transports et du logement (METL). La documentation française*. Paris 74p.

Pistocchi, A., Luzi, L., Napolitano, P., 2002. The use of predictive modelling techniques for optimal exploitation of spatial databases: a case study in landslide hazard mapping with expert system-like methods. *Environmental Geology*, V. 41, p. 765–775.

Soeters, R., & van Westen, C.J., 1996. Slope instability recognition, analysis, and zonation: *Landslides Investigation and Mitigation*,, p. 129–177.

Thiery, Y., Sterlacchini, S., Malet, J.-P., Puissant, A., Maquaire, O., 2004. Strategy to reduce subjectivity in landslide susceptibilityzonation by GIS in complex mountainous environments. In:Toppen, F., Prastacos, P. (Eds.), *Proceedings of AGILE 2004: 7th AGILE Conference on Geographic Information Science*. 29th April–1st May 2004, Heraklion, Greece, pp. 623–634.

Thiery, Y., Malet, J., Sterlacchini, S., Puissant, A., Maquaire, O., 2007. Landslide susceptibility assessment by bivariate methods at large scales: Application to a complex mountainous environment. *Geomorphology*, v. 92, no. 1-2, p. 38-59.

van den Eeckhaut, M.V., Marre, A., Poesen, J., 2009. Comparison of two landslide susceptibility assessments in the Champagne-Ardenne region (France). *Geomorphology*, v.115, no1-2, p. 141-155.

van den Eeckhaut, M., Vanwalleghem, T., Poesen, J., Govers, G., Verstraeten, G., Vandekerckhove, L., 2006. Prediction of landslide susceptibility using rare events logistic regression: a case-study in the Flemish Ardennes (Belgium). *Geomorphology*, v. 76, no. 3-4, p. 392-410.

van Westen, C.J., 2000. The modelling of landslide hazards using GIS. *Surveys in Geophysics*, v. 21, no. 2, p. 241-255.

van Westen, C., Rengers, N., Soeters, R., 2003. Use of Geomorphological Information in Indirect Landslide Susceptibility Assessment. *Natural Hazards*. V. 30 p. 399-419.

van Westen, C., van Asch, T.W.J., Soeters, R., 2006. Landslide hazard and risk zonation-why is it so difficult? *Bulletin of engineering geology and the environment*, V. 65, p. 167-184.

van Westen, C., Castellanos, S., Kuriakose, S.L., 2008. Spatial data for landslide susceptibility, hazard and vulnerability assessment: An overview. *Engineering Geology* v. 102 p. 112-131.