

**Cartographie de la susceptibilité aux glissements de terrain lents dans le bassin tertiaire du Puy-en-Velay : apport des combinaisons de méthodes**

Alexandre Poiraud

► **To cite this version:**

Alexandre Poiraud. Cartographie de la susceptibilité aux glissements de terrain lents dans le bassin tertiaire du Puy-en-Velay : apport des combinaisons de méthodes. Journées Aléa Gravitaire, Sep 2013, Grenoble, France. hal-01088297

**HAL Id: hal-01088297**

**<https://hal.archives-ouvertes.fr/hal-01088297>**

Submitted on 27 Nov 2014

**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.

# Cartographie de la susceptibilité aux glissements de terrain lents dans le bassin tertiaire du Puy-en-Velay : apport des combinaisons de méthodes

A. Poiraud

Chercheur associé, GEOLAB UMR-CNRS 6042, Clermont-Ferrand, France

RESUME : La modélisation de la susceptibilité des mouvements gravitaires est une voie de recherche cherchant à réduire la subjectivité des approches expertes et à fournir des cadres méthodologiques reproductibles pour la cartographie de l'aléa. Depuis une vingtaine d'années maintenant, l'apport des SIG et des méthodes probabilistes a largement nourri la littérature scientifique et les comparaisons de méthodes ont contribué pour une large part à ces apports scientifiques. Une voie encore peu explorée vise à combiner les résultats obtenus par plusieurs méthodes afin de limiter les artefacts de chacune des approches. Cette procédure permet en outre de quantifier l'incertitude de classement des pixels par classe de susceptibilité. Sur le bassin du Puy, un ensemble de 5 méthodes de modélisation de la susceptibilité ont été combinées en une carte de susceptibilité/certitude permettant de quantifier le niveau de susceptibilité et la certitude de classement par pixel. Une validation ROC ainsi qu'une validation terrain permettent de considérer le modèle final comme acceptable. Le modèle permet enfin de pointer les zones de forte susceptibilité mais de forte incertitude sur lesquelles des investigations complémentaires (géotechnique) devront être menées afin de parfaire la cartographie de la susceptibilité pour une possible utilisation dans un PPR.

## 1 INTRODUCTION

La cartographie de la susceptibilité aux glissements de terrain a connu un fort développement méthodologique depuis la fin des années 90 et la démocratisation des SIG (Guzzetti et al., 1999; Chacon et al., 2006). De nombreuses méthodes et algorithmes ont fait l'objet d'expérimentations sur des terrains différents et avec des types de phénomènes gravitaires variés. Ces méthodes de cartographie de la susceptibilité s'appuient essentiellement sur le croisement d'un inventaire de phénomènes gravitaires avec des facteurs de prédispositions locaux afin de dégager les relations entre la présence de ces facteurs et les phénomènes. Ces relations sont ensuite généralisées afin de produire une carte de probabilité de présence future de phénomènes gravitaires.

Après plus de 20 ans de recherche, des standards de procédure commencent à être publiés (Van-Westen, 2004; Begueria, 2006; Guzzetti et al., 2006; Corominas & Moya, 2008; Van-Westen et al., 2008) et la communauté a le matériel nécessaire pour critiquer ces méthodes. Une des premières difficultés rencontrées est qu'il n'existe pas de méthode plus performante que les autres et que leur efficacité dépend des contextes locaux et des données d'entrée. Cependant, on peut tout de même noter une certaine robustesse des méthodes bivariées ou des régressions logistiques qui sont actuellement très utilisées par la communauté (Thiery, 2007). Parmi les critiques ou problèmes demeurant, certains auteurs font remarquer la persistance d'incertitudes récurrentes dans les classes de susceptibilité intermédiaires (Guzzetti *et al.*, 2006; Van-den-Eeckhaut *et al.*, 2009). En effet, si certaines méthodologies peuvent produire des résultats très bien ajustés aux données d'entrée, elles sont cependant moins bonnes pour ce qui est de la prédiction (cas des réseaux de neurones dans l'étude de Rossi et al. (2010).

Au final, deux problèmes sont posés dans cette contribution : i) le problème du choix de la méthode la mieux adaptée au contexte local et aux données d'entrée et ii) la quantification de l'incertitude des classes intermédiaires de susceptibilité afin de réajuster a posteriori le modèle.

## 2 ZONE D'ÉTUDE ET MÉTHODOLOGIE

### 2.1 Zone d'étude

Le bassin du Puy-en-Velay se situe dans le sud-est du Massif central (45°N, 3.5°E, Fig. 1). C'est un bassin tertiaire lié à l'extension oligocène du Rift Ouest Européen, dont il constitue un des principaux prolongements méridionaux (Michon, 2000). Il est encadré à l'ouest par le massif basaltique du Devès d'âge plio-pléistocène, au sud et à l'est par le vaste massif volcanique du Mezenc-Meygal d'âge miocène à pliocène et délimité au nord par le horst granitique de Chaspinhac (Mergoïl & Boivin, 1993). Ses terrains sont essentiellement constitués de matériaux argilo-sableux mécaniquement peu résistants dans les parties périphériques du bassin et d'illites éo-oligocène plus raides dans la partie centrale (Poiraud, 2012). Il est traversé du sud au nord par la Loire qui contrôle l'encaissement de l'ensemble du réseau hydrographique. Enfin, le système de pente est structuré par des phénomènes d'inversion de relief individualisant des mésas basaltiques, des pointements volcaniques (dykes, sills et necks) et des corniches bordières de plateau volcanique. L'ensemble forme un relief hétérogène à forte rugosité.

L'inventaire des phénomènes gravitaires permet en outre de dénombrer un nombre conséquent de glissements lents de type fluage ou solifluxion qui sont à l'origine d'importantes déformations du terrain et de désagréments sur les infrastructures et les habitations. Ces glissements lents peuvent être superficiels (1 à 2m) ou plus profonds (5m). Leur vitesse reste de l'ordre de quelques cm/an et leur activité est liée à l'occurrence de pluies violentes de plus de 100mm/48h (Poiraud, 2012).

Le travail de modélisation a porté plus spécifiquement sur la partie sud-est du bassin (Fig. 1) qui est la plus affectée par ces glissements lents et dont le contexte lithostructural et climatique semble le plus favorable à l'occurrence de ces phénomènes.

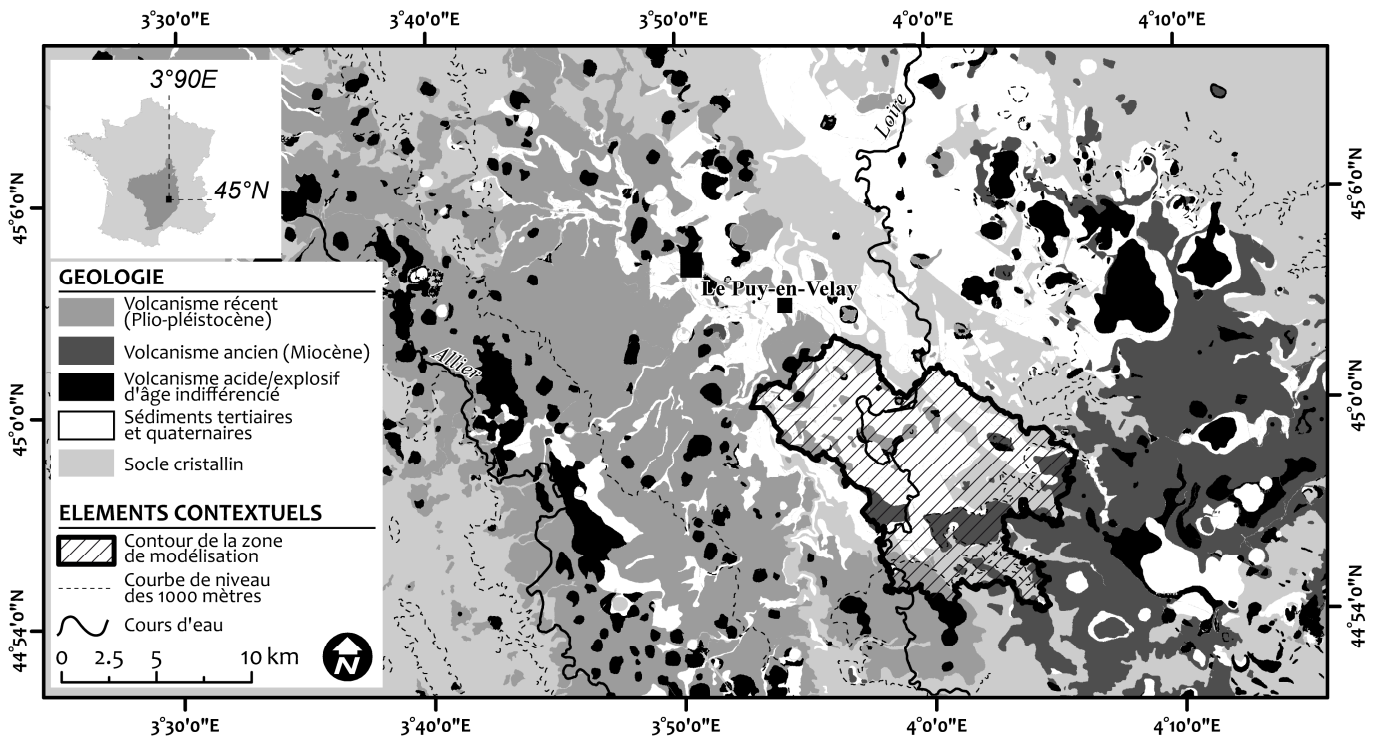


Figure 1. Carte géologique simplifiée du bassin du Puy et ses environs et localisation de la zone d'étude

### 2.2 Méthodologie

L'inventaire des phénomènes gravitaires s'est fait par photo-interprétation avec une vérification « terrain » de tous les polygones cartographiés. Certains glissements représentatifs ont été investigués en détail afin de préciser la typologie ainsi que les processus sous-jacents (suivi cinétique, sondage, géophysique, analyse mécanique des matériaux). Au total, près de 200 phénomènes ont été inventoriés sur un peu moins de 400km<sup>2</sup> mais avec une répartition très inégale des densités.

Le processus de modélisation de la susceptibilité a porté uniquement sur les 200km<sup>2</sup> de la zone sud-est (Fig. 1) et s'est déroulé en 4 étapes (Fig. 2) :

- production d'un jeu de données pour la modélisation et choix des données pertinentes,
- production de 5 cartes de susceptibilité avec 5 méthodes différentes,
- validation de chacune des cartes et estimation de la valeur ROC
- combinaison des 5 cartes en une carte de susceptibilité-certitude et validation ROC

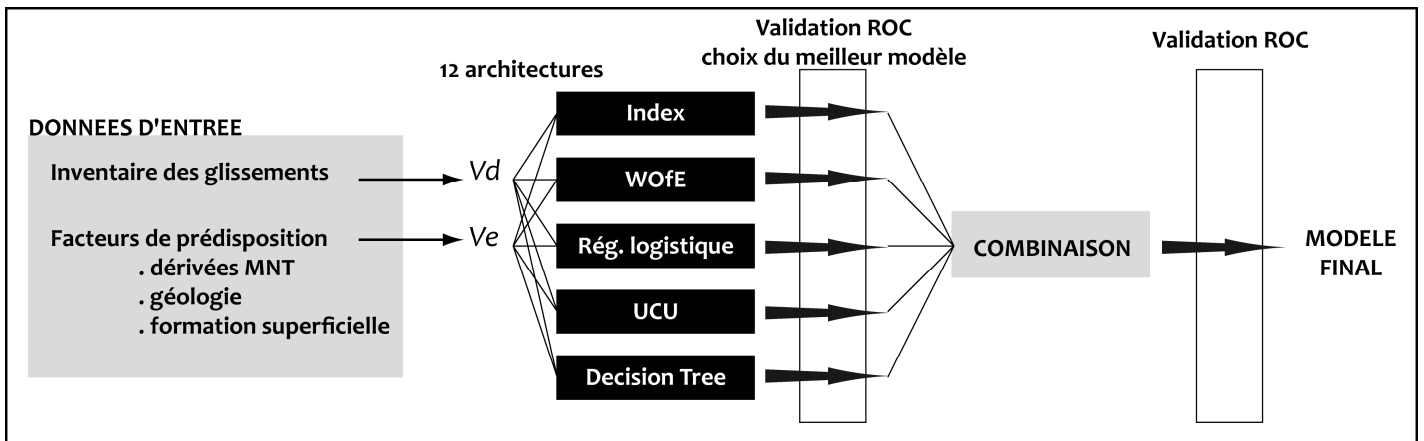


Figure 2. Procédure de modélisation simplifiée.  $Vd$ , variable dépendante ;  $Ve$ , variable explicative

La variable dépendante est composée de 2 sous-groupes choisis aléatoirement dans la population des glissements lents de type fluage : i) un groupe d'entraînement (70% de la population, soit 1456 pixels) et ii) un groupe de validation (30% de la population soit 624 pixels). Les variables explicatives ont été testées afin de mesurer les colinéarités et leur poids explicatifs (relation statistique avec la présence des glissements). Un lot de 8 variables explicatives non colinéaires et ayant suffisamment de poids a été choisi pour la construction des modèles (géologie, formation superficielle, pente, orientation, distance aux thalwegs, TWI, Indice de position topographique et la variable « matériau » qui est une variable fusionnée de la géologie et des formations superficielles). Pour chacune des méthodes, 4 architectures principales composées de 3 sous-architectures de données d'entrée ont été testées (soit 12 architectures au total, Tableau 1). La maille de calcul élémentaire est un pixel de 15m de côté.

Tableau 1. Composition des architectures de données utilisées pour les modélisations

Architecture du modèle	Variables explicatives utilisées
ARCHI1	géologie + pente + orientation
ARCHI2	FS + pente + orientation
ARCHI3	matériau + pente + orientation
ARCHI4	géologie + pente + orientation + distance thalweg
ARCHI5	FS + pente + orientation + distance thalweg
ARCHI6	matériau + pente + orientation + distance thalweg
ARCHI7	géologie + pente + orientation + distance thalweg + TPI
ARCHI8	FS + pente + orientation + distance thalweg + TPI
ARCHI9	matériau + pente + orientation + distance thalweg + TPI
ARCHI10	géologie + pente + orientation + distance thalweg + TPI + TWI
ARCHI11	FS + pente + orientation + distance thalweg + TPI + TWI
ARCHI12	matériau + pente + orientation + distance thalweg + TPI + TWI

Les méthodes ont été appliquées de manière « standard » en suivant les recommandations des auteurs principaux. La méthode des index s'est appuyée sur la pratique du *Likelihood Ratio* (Chung & Fabbri, 2008). L'utilisation de la méthode bivariée de la théorie de l'évidence (WofE) s'est appuyée sur les indications de Bonham-Carter (1994) en utilisant le module ArcSDM (Sawatzky et al., 2004). La méthode multivariée de régression logistique a été implémentée avec SPSS pour obtenir le LOGIT, lequel a été transformé en Odds puis en probabilité (Bai et al., 2010). L'équation obtenue a été appliquée dans ArcGIS pour obtenir les cartes de susceptibilité. L'approche par *Data Mining* s'est appuyée sur la méthode des Arbres de Décision (*Decision Tree*) implémentée également avec SPSS (Nefeslioglu et al., 2010) afin de générer les règles de l'arbre et les appliquer ensuite dans ArcGIS afin de produire la carte de susceptibilité. Enfin, nous avons détourné la méthode des Unités de Conditions Uniques qui sont classiquement utilisées comme unités de calcul (Chung et

al., 1995; Remondo et al., 2003). Dans notre travail de modélisation, chaque UCU est considéré comme une entité caractérisée par la fusion des variables et pour laquelle on calcule une densité de glissement. Par ce biais, nous obtenons une carte de probabilité exprimant la sensibilité de chaque UCU aux instabilités.

Pour chacune des méthodes, nous obtenons 12 modèles (correspondant aux 12 architectures) et 48 modèles pour les arbres de décisions (12 architectures et 4 algorithmes de classifications). Pour chacune des méthodes, nous avons évalué sa capacité de prédiction par l'établissement d'une courbe ROC à partir des pixels du groupe de validation (pixels non utilisés pour la production des modèles) et retenu le meilleur modèle. Nous obtenons 5 modèles finaux, 1 modèle par méthode.

Chacun des 5 modèles a été discrétisé en 4 classes en fonction du rapport de densité de glissement de chacun des centiles sur la densité moyenne de la zone d'étude. Les 5 modèles discrétisés ont ensuite été combinés en calculant la médiane des valeurs de classe pour chaque pixel ainsi que la variété (nombre de classes différentes) permettant d'évaluer la volatilité de classement pour chacun des pixels.

### 3 PRINCIPAUX RESULTATS

#### 3.1 Performance des méthodes et meilleur modèle par méthode

Les valeurs AUC obtenues sont dans l'ensemble satisfaisantes (Fig. 3).

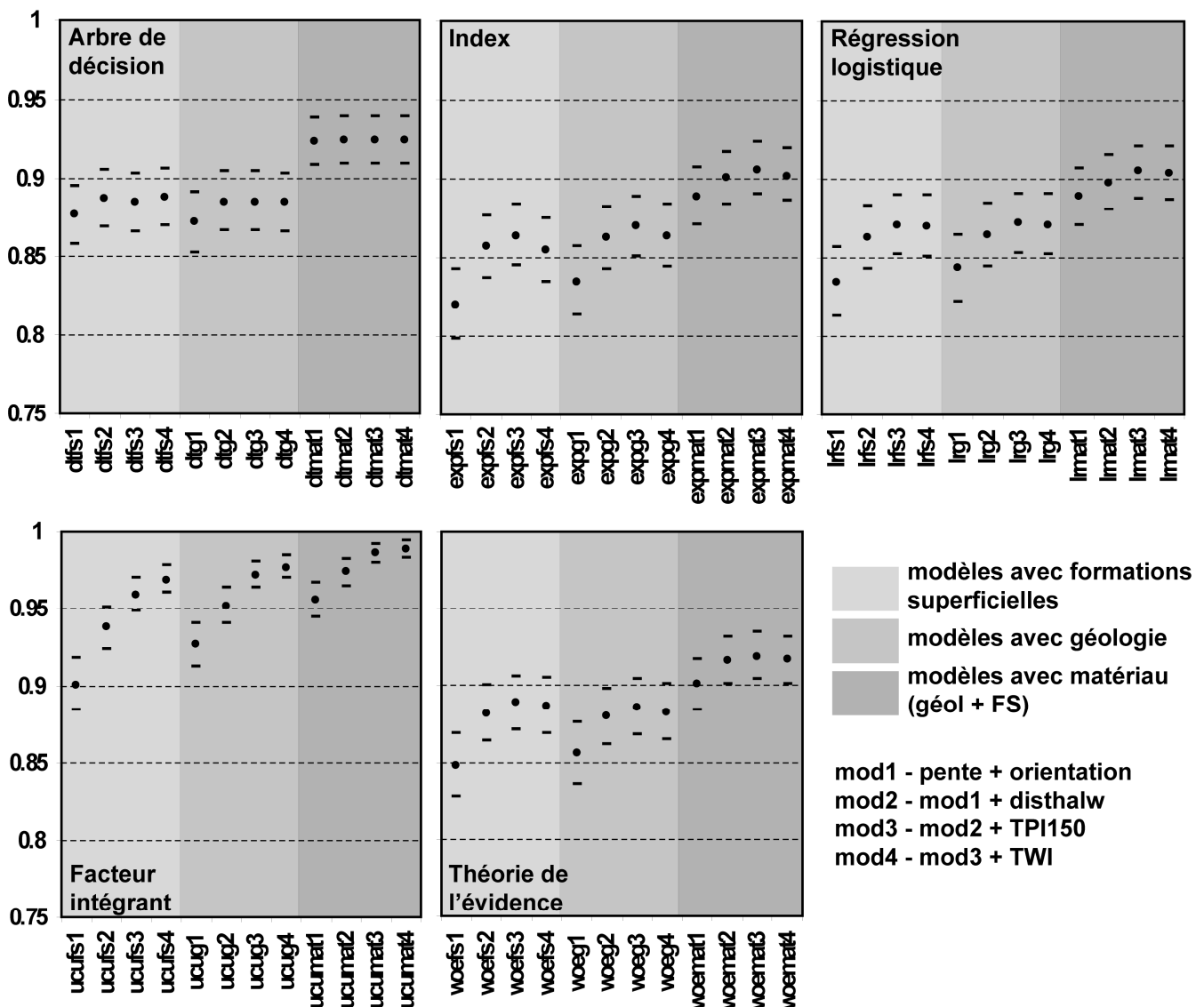


Figure 3. Valeurs d'AUC (Area Under Curve) obtenues par modèle avec le groupe de validation.

Les meilleurs résultats sont toujours obtenus avec la variable intégrée « matériau », celle-ci prenant en compte à la fois la géologie et les formations superficielles. L'apport de cette variable intégrée peut s'expliquer en partie par la plus grande richesse d'information qu'elle contient, mais aussi par la préservation

des interactions entre les variables « géologie » et « formations superficielles ». L'architecture la plus optimale est l'architecture n°9 (Tableau 1) dans la majorité des cas, l'ajout de la variable TWI ayant tendance à faire baisser les performances statistiques du modèle. Nous avons retenu l'architecture 9 pour l'ensemble des méthodes comme étant la plus performante hormis pour les Arbres de Décision ou l'architecture 8 (formation superficielle au lieu de matériau) a des résultats plus cohérents avec la perception « terrain ».

### 3.2 Comparaison des meilleurs modèles

Les 5 modèles retenus convergent dans l'ensemble mais diffèrent dans le détail (fig.4.1). Les valeurs de *kappa* sont très basses pour le modèle issu des Unités de Condition Unique comparées aux autres modèles, du au fait que le résultat obtenu est très bien ajusté aux données d'entrée et a tendance à fournir un résultat binaire. L'arbre de décision diffère également des autres modèles notamment dans la classe de forte susceptibilité. Les modèles probabilistes d'inde, bivarié ou multivarié se ressemblent plus avec des valeurs de *kappa* et de V de Cramer > 0.58. Le calcul des indices de *kappa* par classes et par modèle montre la très grande variabilité de classement des classes intermédiaires (2 et 3) et la plus grande stabilité relative des classes 1 et 4 (Fig. 4.2).

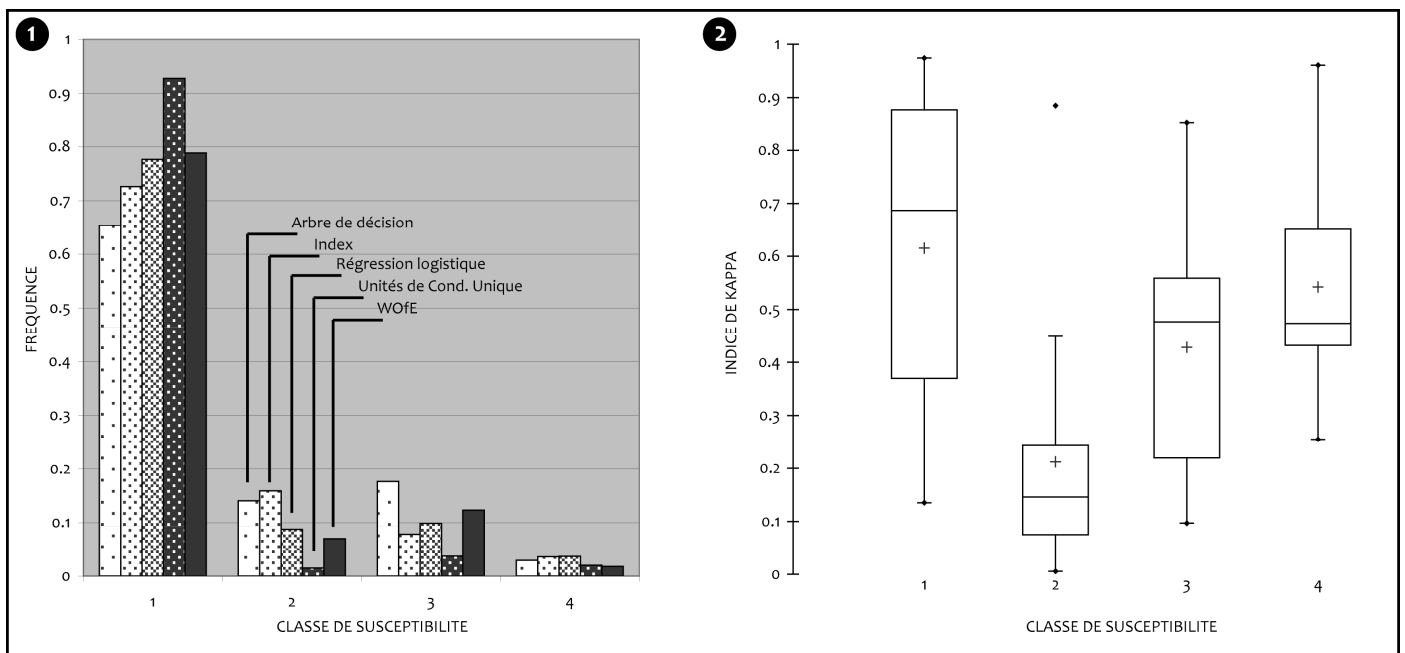


Figure 4. Comparaison des modèles. 1) Fréquences des pixels par classe ; 2) Box-Plot des indices de Kappa par classe

Finalement, chaque modèle propose des classifications acceptables d'un point de vue statistique mais avec des qualités et des défauts de classement d'un point de vue expert. Si les modèles bivarié, multivarié et la méthode des index proposent des classements qui se rapprochent le plus d'un classement expert, les arbres de décision généralisent plus dans les classes intermédiaires ce qui peut apporter une finesse accrue dans ces classes. Le modèle UCU est très binaire mais très performant dans la prédiction de la classe de forte susceptibilité ce qui en fait un modèle intéressant.

### 3.3 Combinaison

Chaque modèle ayant ses arguments favorables, le choix d'un modèle unique est délicat. Felicísimo et al. (2013) avaient pointé cette difficulté et avait proposé la combinaison des modèles en une carte de « consensus » traduisant la moyenne arithmétique des classements. Notre combinaison s'est faite en utilisant la médiane et en calculant la variété ce qui nous donne une idée de la volatilité par classe. La valeur d'AUC atteint 0.92 ce qui est convenable et plus de 56% des pixels sont classés en susceptibilité 1 avec une certitude >80% (5 classements identiques sur 5, Tab. 2). Finalement, 11.1% de la surface totale sont classés avec une certitude inférieure à 60% et demanderaient des investigations complémentaires pour lever les incertitudes. Les zones problématiques sont celles ayant une susceptibilité moyenne à forte (classes 3 et 4) et une certitude faible inférieure à 60%. Ces zones représentent 3.8% de la surface totale.

Tableau 2. Fréquence de pixels par classe de susceptibilité et de certitude

	Susceptibilité				
	S1	S2	S3	S4	
Certitude	<40%	0.1%	0.2%		
	40-60%	1.5%	5.7%	0.2%	
	60-80%	21.2%	3.2%	5.2%	1.6%
	>80%	56.5%	0.1%	0.7%	0.5%

#### 4 CONCLUSIONS

Dans l'esprit des travaux de Felicísimo et al. (2013), notre travail propose l'utilisation de 5 méthodes de modélisation de la susceptibilité aux glissements de terrain puis la combinaison des meilleurs modèles obtenus en une carte de consensus. Nos résultats viennent renforcer d'une part les observations d'une plus grande variabilité et incertitude au sein des classes intermédiaires et d'autre part la difficulté de choisir un modèle optimal, certains étant bon d'un point de vue statistique, d'autres meilleurs d'un point de vue expert. La combinaison par la médiane et le calcul de la variété par pixels offre une solution à ces problèmes en fusionnant les résultats des modèles (lissage des résultats) ce qui évite de faire un « mauvais choix » et en calculant la certitude de classement par pixel ce qui permet de juger d'une part de la convergence des modèles et d'autre part de la qualité finale de la carte obtenue.

En parallèle, l'utilisation d'une variable combinée (variable « matériau », fusionnant la géologie et les formations superficielles) rend les modèles plus performants en conservant les interactions existantes entre les 2 variables initiales.

#### 5 REFERENCES

- Bai, S. B., J. Wang, G. N. Lü, P. G. Zhou, S. S. Hou and S. N. Xu 2010. GIS-based logistic regression for landslide susceptibility mapping of the Zhongxian segment in the Three Gorges area, China. *Geomorphology*, 115: 23-31.
- Beguieria, S. 2006 - Validation and Evaluation of Predictive Models in Hazard Assessment and Risk Management. *Natural Hazards*, 37: 315-329.
- Bonham-Carter, G. F. 1994. *Geographic information systems for geoscientists - modelling with GIS*. Pergamon, 398 p.p.
- Chacon, J., C. Irigaray, T. Fernandez and R. El-Hamdouni 2006. Engineering geology maps: landslides and geographical information systems. *Bulletin of Engineering Geology Environment*, 65: 341-411.
- Chung, C. J. and A. G. Fabbri 2008. Predicting landslides for risk analysis — Spatial models tested by a cross-validation technique. *Geomorphology*, 94: 438-452.
- Chung, C. J., A. G. Fabbri and C. J. Van-Westen 1995. Multivariate regression analysis for landslide hazard zonation. In A. Carrara & F. Guzzetti (eds), *Geographical Information Systems in Assessing Natural Hazards*. Netherlands, Kluwer Academic Publishers, 107-133
- Corominas, J. and J. Moya 2008. A review of assessing landslide frequency for hazard zoning purposes. *Engineering Geology*, 102: 193-213.
- Felicísimo, A., A. Cuartero, J. Remondo and E. Quiros 2013. Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: a comparative study. *Landslides*, 10 (2): 175-189.
- Guzzetti, F., A. Carrara, M. Cardinali and P. Reichenbach 1999. Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. *Geomorphology*, 31: 181-216.
- Guzzetti, F., P. Reichenbach, F. Ardizzone, M. Cardinali and M. Galli 2006. Estimating the quality of landslide susceptibility models. *Geomorphology*, 81: 166-184.
- Mergoïl, J. and P. Boivin 1993. Le Velay, son volcanisme et les formations associées : notice à la carte au 1/100000. *Géologie de la France*, 3, 96p.
- Michon, L. 2000. *Dynamique de l'extension continentale - Application au Rift Ouest-Européen par l'étude de la province du Massif Central*. Thèse de doctorat, Université Blaise Pascal, Clermont-Ferrand II, 266.p.
- Nefeslioglu, H. A., E. Sezer, C. Gokceoglu, A. S. Bozkir and T. Y. Duman 2010. Assessment of Landslide Susceptibility by Decision Trees in the Metropolitan Area of Istanbul, Turkey. *Mathematical Problems in Engineering*.
- Poiraud, A. 2012. *Les glissements de terrain dans le bassin tertiaire volcanisé du Puy-en-Velay Massif central, France : caractérisation, facteurs de contrôle et cartographie de l'aléa*. Thèse de doctorat, Département de Géographie, Université Blaise Pascal, Clermont-Ferrand, 563.p.

- Remondo, J., A. Gonzales-Diez, J. R. Diaz-De-Teran and A. Cendrero 2003. Landslide Susceptibility Models Utilising Spatial Data Analysis Techniques. A Case Study from the Lower Deba Valley, Guipúzcoa Spain. *Natural Hazards*, 30: 267-279.
- Rossi, M., F. Guzzetti, P. Reichenbach, A. C. Mondini and S. Peruccacci 2010. Optimal landslide susceptibility zonation based on multiple forecasts. *Geomorphology*, 114: 129-142.
- Sawatzky, D. L., G. L. Raines, G. F. Bonham-Carter and C. G. Looney 2004. ArcSDM2: ArcMAP extension for spatial data modelling using weights of evidence, logistic regression, fuzzy logic and neural network analysis - <http://ntserv.gis.nrcan.gc.ca/sdm/>.
- Thiery, Y. 2007. *Susceptibilité du bassin de Barcelonnette Alpes du sud, France aux "mouvements de versant" : cartographie morphodynamique, analyse spatiale et modélisation probabiliste*. Université de Caen-basse Normandie, thèse de doctorat, 443 p.p.
- Van-den-Eeckhaut, M., P. Reichenbach, F. Guzzetti, M. Rossi and J. Poesen 2009. Combined landslide inventory and susceptibility assessment based on different mapping units: an example from the Flemish Ardennes, Belgium. *Natural Hazards and Earth System Sciences*, 9: 507-521.
- Van-Westen, C. J. 2004. Geo-Information tools for landslide risk assessment: an overview of recent developments. In W. Lacerda, M. Ehrlich, S. Fontoura, A. Sayao (eds), *Landslides Evaluation & Stabilization*, Proceedings of the 9th International Symposium on Landslides, Rio de Janeiro, Brasil, 28th June-2nd July, 39-56
- Van-Westen, C. J., E. Castellanos and S. L. Kuriakose 2008. Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. *Engineering Geology*, 102: 112-131.