
Variables dominantes dans l'organisation d'un paysage

Méthodologie multiéchelle dans un contexte SIG

Pierre-Olivier Mazagol

*Centre de Recherche sur l'Environnement et l'Aménagement
UMR CNRS 5600
Université Jean Monnet
6 rue Basse des Rives
42023 Saint-Etienne cedex2
pierre.olivier.mazagol@unvi-st-etienne.fr*

RÉSUMÉ. Cherchant à mettre en évidence des variables participant à l'organisation spatiale d'un paysage, nous nous sommes rapidement trouvés confrontés à différents problèmes. Tous sont liés à l'outil d'analyse spatiale que nous avons choisi (Pegase) qui possède son propre format de bases de données. Nous présentons ici la méthodologie que nous avons mis en place afin d'intégrer cet outil dans un contexte SIG. Nous montrons comment son application, selon une approche multiéchelle, permet de mettre en évidence des variables dominantes dans l'organisation d'un paysage.

ABSTRACT. While trying to find out the variables that participate in landscape organization, we were soon confronted to a number of problems. These problems are mainly due to the spatial analysis tool we have chosen (Pegase), which has his own data format. We present in this article our methodology that allows us to use this tool in a GIS context. We show the interest of this application, in a multiscale approach, in order to determine « dominant variable » in landscape organization.

MOTS-CLÉS : paysage, organisation spatiale, analyse multiéchelle, néguentropie, hétérogénéité, variables dominantes, Pegase, SIG

KEYWORDS : landscape, spatial organization, multiscale analysis, neguentropy, heterogeneity, dominant variables, Pegase, GIS

1. Introduction

Parler d'étude, de modélisation, d'évaluation de paysages nécessite donc de définir ce que l'on entend par paysage. Dans notre approche, le paysage se divise en deux unités élémentaires : un support physique et sa couverture (Delpoux 1972) impliquant de nombreuses variables écologiques. Le support physique renvoie aux caractéristiques géologiques, orographiques, géomorphologiques, édaphiques et la couverture à des variables biologiques incluant l'action humaine. Tous deux sont le siège d'échanges. Le paysage apparaît donc comme un système organisé, propice à l'étude de flux, interactions, relations. Il est un objet géographique spatialisé, physique, anthropisé, socio-économique, mais également un objet écologique.

Ainsi, trois approches du paysage peuvent être envisagées (Farina 1998) :

- géobotanique : centrée sur les facteurs environnementaux abiotiques et biotiques dont les facteurs anthropiques ;
- animale : centrée sur la perception de l'environnement de l'organisme vivant considéré ;
- humaine : centrée sur la perception par l'Homme, le paysage étant formé par des entités fonctionnelles qui ont un sens pour la vie de l'Homme.

Ces trois approches sont complémentaires : pour une gestion durable du paysage, la vision centrée sur l'homme et la satisfaction de ses besoins est incontournable, autant que la prise en compte des facteurs et processus abiotiques et biotiques qui structurent le paysage et des besoins des organismes vivants. Le paysage est un espace hétérogène, cohérent et dynamique résultant de la conjonction de facteurs écologiques, sociaux, culturels et économiques. C'est un système constitué d'une mosaïque d'écosystèmes en interaction intégrant des processus naturels (biophysique, énergétique...) et anthropiques.

Dans notre recherche de variables contribuant à l'organisation d'un paysage, et par extension de tout système spatial, nous avons utilisé la procédure PEGASE (Partition d'un Ensemble Géographique pour l'Analyse Spatiale Ecologique). Basée sur les principes de la Théorie de l'Information, cette méthode développée par Michel Phipps (Phipps 1981), utilise un algorithme qui permet la mise en évidence, la hiérarchisation et la quantification des facteurs explicatifs de l'organisation de l'espace. Elle a déjà été éprouvée dans de nombreuses études portant sur des problématiques variées (Phipps et al. 1986b, 1986a; Dumanski et al. 1987; Davis et Dozier 1990; Casado et al. 1997; Lundström-Gilliéron 1999; Ernout et al. 2003; Lundström-Gilliéron et Schlaepfer 2003; Mazagol 2003; Phipps 2003; Mazagol 2006; Mazagol et Porteret 2006b, 2006a; Mazagol 2007). S'appuyant sur la théorie des systèmes discrets ou distribués et la théorie de l'information, elle cherche à mettre en œuvre le concept d'ordre - ou organisation - dans un paysage et à en donner une mesure. L'algorithme permet ainsi d'estimer la contribution de variables (explicatives) à l'organisation d'un espace (caractérisée par exemple par l'occupation du sol (variable cible). La méthode permet de mesurer l'information

contenue dans un paysage (complexité), correspondant également à l'information perçue par l'observateur, et d'apprécier les échanges d'information entre écosystèmes (messages allant d'une unité spatiale à une autre) ou du milieu abiotique au milieu biotique. Cette notion d'information permet une étude quantitative et globale sans nécessiter une connaissance fine des mécanismes de fonctionnement du paysage. Ainsi, l'outil permet d'aborder des questions sous-jacentes à la problématique de l'organisation : prédictibilité des états locaux, reconnaissance des contraintes pesant sur leur distribution spatiale et, indirectement, mise en évidence de relations fonctionnelles liant la variable-cible aux autres descripteurs du système (Phipps 2003). Le formalisme, la solidité mathématique de l'algorithme nous a paru présenter un intérêt majeur, induisant le choix de son utilisation dans le cadre de notre recherche.

Parmi les applications existantes, aucune n'a été entreprise dans un contexte multiéchelle. Or, l'importance des études multiéchelles dans le cadre des analyses écologiques de paysage est d'un grand intérêt (Turner et al. 1989; Jelinski et Wu 1996; Wu et al. 2000; Wu et al. 2002). Les structures spatiales apparentes et donc l'hétérogénéité sont fonction de l'échelle de résolution (Turner et Gardner 1990). L'étude d'un système à une échelle donnée peut ne pas faire apparaître des phénomènes qui apparaîtraient à d'autres échelles (Allen et Star 1982). Les variations de l'aire étudiée en modifie également les caractéristiques globales telles que l'hétérogénéité ou la diversité (Baudry 1992). La notion d'échelle spatiale fait donc référence, en générale, à deux mesures fondamentales : l'étendue qui est la taille de l'espace étudié et la résolution qui correspond à la taille du grain, du pixel ou de la cellule. Outre résolution et étendue, nous rajoutons une autre composante à l'échelle. Certaines mesures sont obtenues par voisinage d'unités (en l'occurrence des cellules ou pixels). Les conditions d'application de formules mathématiques ont une influence marquée sur les résultats obtenus en exprimant des phénomènes différents et mettant en évidence plus ou moins précisément les traits du relief (Mazagol 2006).

Mais l'utilisation de Pegase nécessite l'intégration préalable, au sein d'une base de données, d'une variable cible et de variables explicatives structurant potentiellement le système. Nous présenterons rapidement le terrain d'étude puis, de manière détaillée, la méthodologie au sein de laquelle nous avons placé Pegase. Enfin, nous indiquerons succinctement les résultats obtenus, qui ont déjà fait l'objet d'une communication à ThéoQuant 2007 (Mazagol 2007) et ajouterons un nouvel élément propre à valider nos résultats.

2. Terrain d'étude

Moyenne montagne cristalline située dans le quart Nord-Est du Massif Central (Figure 1), les Monts du Forez sont dominés par une zone sommitale de landes, ou Hautes Chaumes, au dessus de la limite supérieure actuelle de la forêt (1200 mètres

environ) (Thébaud 1988). Leur ligne de crêtes s'étend sur plus de 40 kilomètres de long et culmine à 1634 mètres (Baulig 1928). Les formes massives et la composition de roches exclusivement cristallines en font un exemple didactique de horst (Etlicher 1986). Essentiellement granitiques, elles offrent des reliefs assez contrastés, un peu marqués par les glaciations quaternaires (Etlicher 1986). Malgré une altitude modeste, l'étagement s'étale avec une rare vigueur, qu'il s'agisse de celui de la végétation, des sols, des modelés, des activités (Etlicher 1986). Les Hautes Chaumes, paysage de landes, de pelouses et de tourbières s'étendent sur environ 8500 hectares. La végétation, derrière son apparente monotonie, révèle une hétérogénéité majeure au niveau des communautés qui la constituent (Etlicher et al. 1993). Les contraintes écologiques fortes ont entraîné le développement de biotopes diversifiés et spécialisés d'une grande richesse biologique, caractéristiques de l'étage montagnard supérieur et même de l'étage subalpin. (Thébaud 1990). Certains milieux du territoire relèvent ainsi d'habitats d'intérêt communautaire, imbriqués de manière complexe et constituant autant de biotopes pour des espèces animales et végétales rares. Les Hautes Chaumes possèdent également des milieux humides surtout représentés par les tourbières acides. Elles sont soumises à la fois à des contraintes physiques sévères et à une forte pression anthropique. Du point de vue des conditions physiques, on note une dissymétrie accusée, renforcée par la direction méridienne du massif : un versant au caractère océanique marqué s'oppose à un versant oriental au caractère continental (Etlicher et al. 1993). Au Nord et au Sud, on observe des plateaux aux croupes molles séparés par des dépressions humides à faibles pentes. La partie centrale est caractérisée par des crêtes et des zones rocheuses aux pentes prononcées. Cette topographie est favorable à la variation de pentes et d'exposition dont l'importance est capitale à cette altitude (Etlicher et al. 1993). Les précipitations sont comprises en moyenne entre 1200 mm et 2000 mm annuels et présentent un gradient Nord-Sud. La température moyenne annuelle est généralement inférieure à 6°C. La saison hivernale s'étale en effet de septembre à juin et l'enneigement est important (3 à 6 mois). Le relief génère cependant de nombreux microclimats très différenciés. Les vents violents provoquent la mise en place de nombreuses congères. La rigueur des conditions de vie est accentuée par l'altitude et la présence de sols pauvres, acides (pH entre 4 et 5) et peu évolués (rankers d'altitude). La présence de milieux ouverts d'intérêt européen correspond à la partie subalpine de l'étagement mais également à un défrichement de la hêtraie-sapinière sommitale (Janssen 1990). Des indices polliniques d'anthropisation ont été datés à 5000 avant Jésus-Christ (Cubizolle 2005). L'activité pastorale a permis d'une part de conserver les pelouses et les landes sèches gagnées sur la forêt et d'autre part d'accentuer le contraste entre les prairies régulièrement fumées et les secteurs de landes pâturées.

Traditionnellement dévolues aux activités pastorales d'estive, les Hautes Chaumes, depuis les années 50, sont désormais vouées à de nouvelles pratiques agricoles en rupture complète avec les modèles traditionnels (Fel 1962; Couhert 1988). En effet, ce secteur a subi une déprise agricole qui se traduit par un abandon progressif ou une surexploitation des parcelles pâturées, un regain des espèces de

landes et une extension des surfaces forestières par reconquête naturelle ou reboisement volontaire. Cette sous-utilisation pastorale a déclenché une dynamique naturelle permettant à la forêt de « grignoter » les landes, ce qui risque à moyen et long terme de faire disparaître ces milieux ouverts (Jacqueminet et al. 2002). L'exploitation pastorale est déterminante de la nature ligneuse ou herbacées des différentes formations végétales (Thébaud 1990).

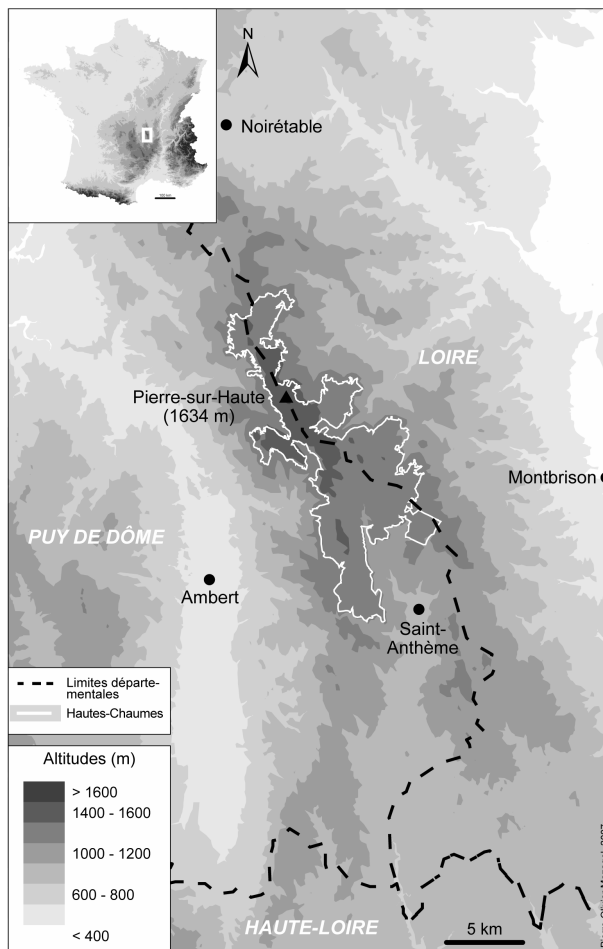


Figure 1. Localisation des Hautes Chaumes des Monts du Forez

Les fumades forésiennes constituent un type de prairies de fauche bien particulier du fait de leur localisation topographique et de leur système d'exploitation pastorale très caractéristique (D'Alverny 1907). Elles sont localisées à proximité des jasseries qui sont des unités d'exploitation servant d'abri pour les bêtes, leurs gardiens, les réserves de nourriture et d'ateliers de fabrication de la fourme avec cave d'affinage et entrepôt à fromage. Ces jasseries sont les sources de l'engrais pour la fertilisation des fumades qui reçoivent, pendant la durée de l'estive, les excréments du troupeau et les eaux de lavage de l'étable (Damon 1972). Régulièrement fertilisées et fauchées, elles font l'objet d'améliorations tant que la jasserie est fréquentée (Etlicher et al. 1993).

Par ailleurs, les Hautes Chaumes constituent une zone touristique très attractive et donc très fréquentée, malgré son éloignement des grands centres urbains, en raison de sa diversité paysagère. Cette pression anthropique conduit à une banalisation biologique et paysagère des biotopes. Cette évolution s'exprime par la diminution de la qualité pastorale des estives induisant la fermeture des paysages, les terres non pâturées étant recolonisées par les ligneux. À l'heure actuelle, les milieux ouverts ne représentent plus que 63% du site. Il en découle une préoccupation des gestionnaires, à la fois pour le maintien des activités et des populations et le développement touristique.

3. Méthodologie

3.1. Échantillonnage de zones

Nous avons sélectionné cinq zones-tests. Si les limites de ces zones n'ont aucune signification scientifique, toutes possèdent une « signature » (distribution et fréquence des types d'occupation du sol) qui leur est propre (Figure 2). Les superficies sont choisies afin que le nombre de cellules soit important aux hautes résolutions et suffisant aux plus faibles. Elles devaient permettre l'augmentation de leur étendue. Les zones 3, 4 et 5 sont localisées dans la partie centrale des Hautes Chaumes.

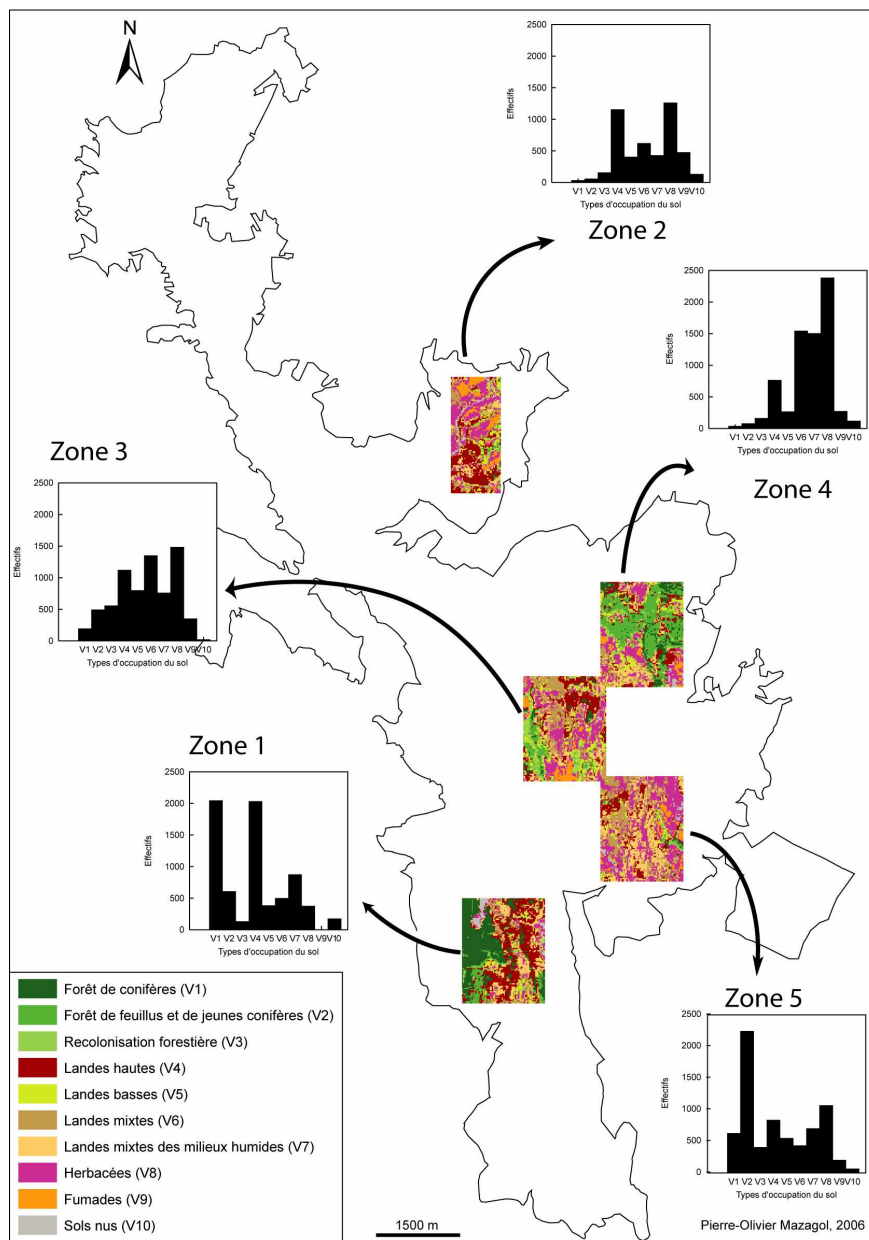


Figure 2. Sélection des zones tests

3.2. Mise en place de la méthodologie

Notre méthodologie s'articule en 5 points principaux (Figure 3) :

- acquisition et production des données au format raster ;
- centralisation pour exportation des données au format Esri ascii grid ;
- intégration de ces données au sein d'une base de données Pegase ;
- traitement des bases de données ;
- analyse des résultats, représentation graphique, récupération des cartes de sortie Pegase.

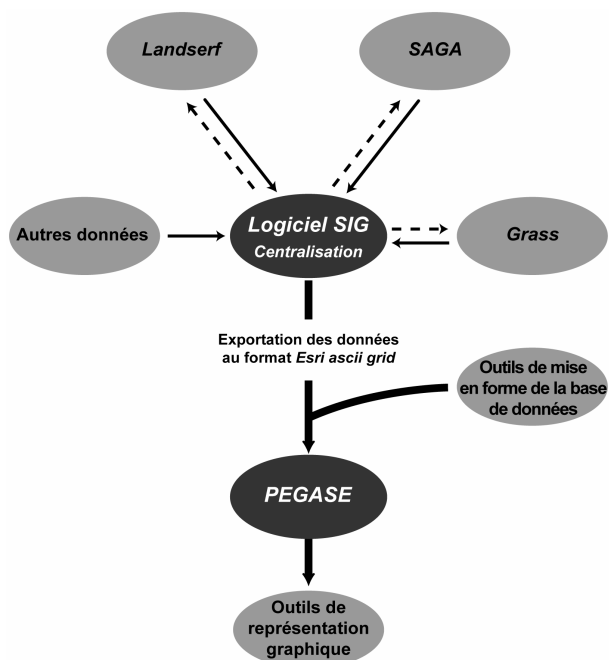


Figure 3. Mise en place de la méthodologie, autour de Pegase, dans un contexte de Système d'Information Géographique

3.2.1. Les données utilisées

Les données utilisées ont plusieurs origines et différents formats (Tableau 1). La principale est l'altitude obtenue grâce au modèle numérique de terrain (MNT) de l'IGN, au format raster. À partir de ce dernier, grâce à l'application de d'algorithme, il est possible d'obtenir un grand nombre de variables, d'autant plus que certaines d'entre elles sont estimées à l'aide de noyaux de pixels de tailles différentes.

	Stratégie de reclassement (Tailles des noyaux de pixels)	Origine
Altitude	50 m	MNT I.G.N.
Exposition	N, NE, E, SE, S, SO, O, NO (3×3, 5×5, 9×9)	Lanserf 2.2
Pente	2,5° (3×3, 5×5, 9×9)	Lanserf 2.2
Courbure verticale	Négative, nulle ou positive (3×3, 5×5, 9×9)	Lanserf 2.2
Courbure horizontale		Lanserf 2.2
Courbure transversale		Lanserf 2.2
Courbure longitudinale		Lanserf 2.2
Courbure maximale		Lanserf 2.2
Courbure minimale		Lanserf 2.2
Courbure moyenne		Lanserf 2.2
Indice ATD		Accumulation, Transit ou Dispersion (3×3, 5×5, 9×9)
Rugosité	4 m ² (3×3, 5×5, 9×9, 20×20 et 30×30)	SIG
Relief local relatif	15 m (3×3, 5×5, 9×9, 20×20 et 30×30)	SIG
Radiation solaire	75 kWh/m ²	SAGA 2.0
Durée d'insolation	150 h	SAGA 2.0
Indice topographique	1 (unité arbitraire)	Grass 6.1
Distance aux tourbières	50 m	Porteret 2001 + SIG
Distance au réseau hydrographique	50 m	BD Carthage + SIG
Distance aux talwegs	50 m	Landserf 2.2 + SIG
Distance aux crêtes	50 m	Landserf 2.2 + SIG
Distance aux chemins	50 m	SIG
Distance aux zones humides	50 m	SIG
Formations superficielles	Gros blocs de surface, arène meuble à blocs, roche subaffleurante, dallage sous-jacent avec arène meuble en surface, gros blocs de basalte en surface, basalte subaffleurant	Ettlicher, 1986
Parcellaire	terrains privés, collectifs	Couhert, 1988 ; Bessenay, 1998
Fréquentation animale	ovine, bovine, mixte ou nulle	Couhert, 1988 ; Bessenay, 1995
Utilisation	Individuelle ou collective par des SICA, coopératives ou groupements	Couhert, 1988 ; Bessenay, 1995
Exploitations des surfaces	terrains non utilisés ou boisés, pâturés et fauchés	Couhert, 1988 ; Bessenay, 1995

Tableau 1. Les différentes variables explicatives sélectionnées

D'autres sont des données vectorielles, lignes et polygones. Elles peuvent être utilisées en l'état (c'est le cas des données surfaciques), ou alors soumises à bufferisation (dans le cas des données linéaires ou ponctuelles) permettant alors de prendre en compte la portée spatiale de ces éléments. Toutes ces données sont choisies en fonction de leur rôle potentiel dans l'organisation du paysage.

Il est cependant nécessaire de mener une réflexion quant à la stratégie utilisée pour le reclassement des données. Dans notre recherche, le reclassement a été effectué sur la totalité du terrain d'étude, afin que dans chaque zone, quel que soit sa localisation et/ou son niveau, une classe ait un code et un seul. Mais une telle stratégie de reclassement présente des limites. En effet, plus l'étendue de la zone est importante, plus la probabilité que l'amplitude de certaines variables soit forte est importante. Dans ce cas, le nombre de classes peut être très élevé, ce qui rend l'exécution de Pegase périlleuse. De plus, dans le cas des variables estimées par noyaux de cellules, la taille de ces derniers influe sur l'amplitude de la valeur calculée. Il nous a donc fallu jongler entre deux impératifs : un nombre de classes raisonnable et un reclassement ayant une signification pour variable considérée. Cette étape est donc subjective et dépend également des contraintes liées aux capacités des différents éléments de la chaîne logicielle.

Pour une analyse critiques de ces données plus poussée, reportez vous à (Mazagol 2006)¹.

L'hétérogénéité des données mises en jeu dans nos bases de données se double de la multiplicité des logiciels et donc des formats utilisés. Cela implique l'utilisation d'un logiciel SIG centralisateur autour duquel on place les autres outils comme applications satellites. Elles sont ensuite exportées au format Ascii Esri Grid, qui permet la construction des bases de données spécifiques au logiciel

3.2.2. *Mise en place des bases de données*

Les bases de données Pegase sont des fichiers « texte » à l'extension « .pgd ». et à structure spécifique. Elles sont constituées en deux parties :

- un en-tête contenant diverses informations et notamment les options qui seront utilisées pour mener à bien la procédure divisive ou les variables présentes dans la base de données ;
- un corps renseignant pour chaque cellule les attributs de chaque variable.

Pour plus d'informations, il est nécessaire de se reporter au manuel d'utilisation du logiciel Pegase (Phipps 2002).

Pour en automatiser la mise en place à partir des fichiers exportés précédemment, nous avons développé deux outils :

1. Chapitre IV : Évaluation critiques des données mobilisées - thèse téléchargeable à l'adresse : <http://mazagol.free.fr/spip.php?article27>

- Asc2Peg (Mazagol et Mayoud 2005) avec lequel nous créons le corps ;
- XLPeg (Mazagol 2005) qui permet la structuration de l'en-tête.

Pour plus d'information sur le fonctionnement de ces deux outils, reportez-vous à Mazagol (2006).

3.2.4. Pegase

Ainsi, l'espace étudié est caractérisé par un système cellulaire. Chaque unité spatiale (chaque cellule) est caractérisée par des descripteurs biotiques, abiotiques ou autres qui lui sont associés. Pegase divise l'ensemble des unités, par paliers, en sous-ensembles significatifs (Figure 4). Ces derniers sont de mieux en mieux définis par les variables explicatives (divisives). Paliers après paliers, Pegase détermine quelle(s) variable(s), parmi les variables explicatives à disposition, permet(tent) de réduire l'entropie initiale au maximum, cette dernière étant estimée selon la formule de Shannon (Phipps 1981; Lundström-Gilliéron 1999; Phipps 2002, 2003; Mazagol 2006). Le paysage est découpé en zones qui rassemblent les unités où cette variable divisive prend la même valeur (Terminal subsets : TSS). Pour chaque TSS, on aura une probabilité différente d'être en présence d'une certaine valeur de la variable cible sous l'influence d'une combinaison particulière des variables explicatives, représentant la réduction maximum du nombre de configurations possibles. On a introduit une contrainte dans le paysage, créé un ordre. En fin de processus, l'entropie du système au maximum, en fonction d'une série de différents critères.

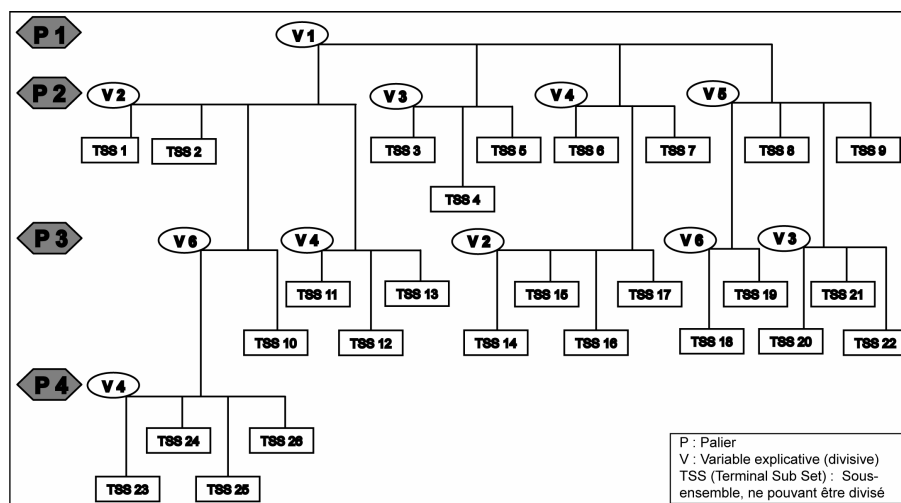


Figure 4. Dendrogramme de la division hiérarchique

La différence entre entropie initiale et entropie finale est la néguentropie. Exprimée en pourcentage, elle donne la redondance. La spécificité correspond au degré de dépendance des classes de la variable cible vis-à-vis des combinaisons de variables explicatives soulevées par l'algorithme. Elle montre dans quelle mesure la présence d'une classe dans un TSS exclut celle des autres classes. Enfin, l'état prédit correspond, pour chaque cellule, à une classe de la variable cible tirée au hasard, par la méthode de Monte-Carlo, en tenant compte de probabilités conditionnelles définies dans l'algorithme. En effet, chaque cellule appartient à un TSS qui indique les degrés d'affinité de ce dernier avec les classes de la variable cible, et auquel est associé une combinaison de variables explicatives. Pegase fournit différentes métriques dont l'entropie initiale, l'entropie finale, la néguentropie, la redondance, la contribution de chaque variable à la néguentropie et la spécificité.

3.2.5. *Représentation graphique des résultats*

Pegase fournit également des représentations graphiques de ses résultats, mais ces dernières ne peuvent pas être exportées autrement que par des captures d'écran de piètre qualité, rendant alors leur exploitation peu aisée. Pour pallier ce problème, nous avons utilisé le logiciel libre Gnuplot 4.0 qui permet, à partir d'informations contenues dans un fichier texte, de tracer des graphiques exportables au format vectoriel, assurant ainsi un meilleur rendu. Autre avantage, il permet de tracer des graphiques de différents types en deux et surtout en trois dimensions, ce qui présente un intérêt dans le cadre de notre travail, au vu la taille de nos bases de données. Ces graphiques seront réalisés à partir des informations présentes dans le fichier de sortie Pegase (*.pgl). Par contre, la cartographie de l'état prédit ne peut être exportée que par capture d'écran.

3.2.6. *Stratégie multiéchelle*

Les zones 3, 4 et 5 sont localisées dans la partie centrale des Hautes Chaumes. Le domaine étant moins étroit, c'est à partir d'elles que l'augmentation de l'étendue des zones est la plus facilement réalisable. Cette dernière se fait selon les niveaux II, III puis IV (Figure 5). Il est à noter que les zones z3nIV, z4nIV et z5nIV sont identiques. Les zones 1 et 2 ne sont donc étudiées qu'à une seule étendue.

Si l'on considère leur nature, les données en notre possession peuvent être classées dans trois groupes :

- celles au format raster, qui possèdent donc une résolution d'origine (l'occupation du sol : 20 m) ;
- celles au format raster dont la résolution dépende d'une autre donnée source (les variables dérivées du MNT) ;
- celles au format vecteur dont la résolution dépendront du choix de l'utilisateur lors de leur conversion au format raster.

Sur ce dernier point, il est important de préciser que cette transformation ne peut se faire dans n'importe quelles conditions. En effet, les données vectorielles ne sont pas utilisables à toutes les résolutions, du fait d'une diminution de la qualité de l'information. Ainsi, le choix dépend également de l'étendue de la zone d'étude. Nous avons opté pour une échelle de résolution allant de 20 à 100 m, par pas de 10 m, considérant que :

- au-delà de 100 m, la perte d'information serait trop importante, notamment pour les zones de niveau I ;
- en-deçà de 20 m, le gain d'information ne serait pas significatif.

Pour une analyse des effets de changement d'échelle, reportez vous à (Mazagol 2006).

Nous construisons donc 108 bases de données Pegase : une par étendue et par résolution considérée.

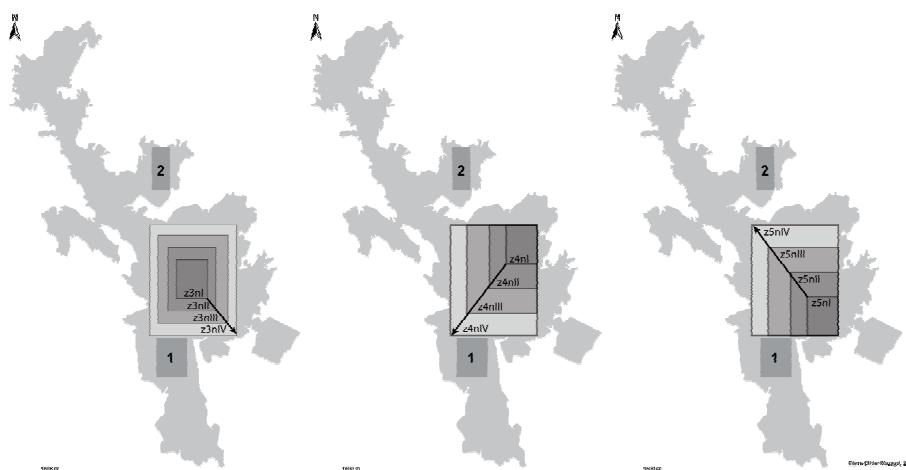


Figure 5. Variation de l'étendue à partir des zones 3, 4 et 5

4. Résultats

Les résultats obtenus pour la zone 3 ayant déjà été publiés (Mazagol 2007), nous présentons ici ceux de la zone 4. Nous irons donc à l'essentiel et ajouterons un

résultat supplémentaire. La néguentropie correspond à la réduction de l'entropie à l'issue de l'application de l'algorithme. Son expression en pourcentage est la redondance. Pour la zone 4, cette dernière est proche de 60% à la résolution 20 m quelle que soit l'étendue. Ainsi, 60 % de l'organisation du paysage est soulevée par les variables explicatives sélectionnées. Ce sont elles qui exercent une contrainte sur le système spatiale. Globalement, ce sont les étendues les plus élevées (III et IV) qui permettent d'obtenir la meilleure redondance et ce, à chaque résolution, l'aspect des courbes étant sensiblement le même pour ces niveaux. Dès une résolution 50 m, la courbe du niveau I a un comportement erratique caractéristique d'un nombre insuffisant de cellules pour valider les résultats du processus. Un nombre élevé de cellules (fortes étendue et résolution) permet de mieux cerner le système ce qui induit une redondance plus importante. Il en va de même pour la courbe de l'étendue de niveau II, à partir de la résolution 70 m.

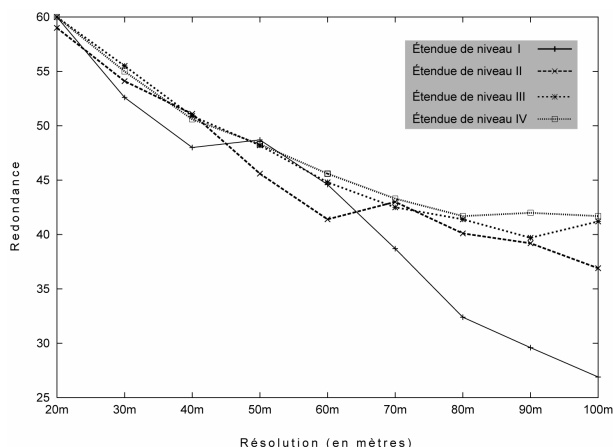


Figure 6. Évolution de la redondance en fonction de la résolution, aux différentes étendues considérées (zone 4)

Ces constatations sont valables quelle que soit la zone considérée (± 10 m de résolution). Ainsi, nous avons choisi de ne pas tenir compte des résultats obtenus aux résolutions supérieures à 50 m, pour l'étendue de niveau I, et 80 m, pour l'étendue de niveau II. Pegase fournit la contribution (en pourcentage) de chaque variable à la néguentropie. Il s'agit donc de la part qu'elles détiennent dans l'organisation du système. Ces contributions peuvent être infimes voire nulles. De plus en fonction, de la résolution considérée ou de l'étendue, elles peuvent varier. Pour prendre en compte cette constatation, donc l'influence de l'échelle, nous avons

défini deux règles de sélection des variables explicatives ont été mises en place :

- la première règle (issue d'une réflexion sur les résultats de nos différents traitements) stipule que la contribution d'une variable à la néguentropie est significative si elle est supérieure à 5 %. Certaines variables sont alors écartées ;
- la deuxième règle impose qu'une variable est dominante, si elle contribue de manière significative à plus de la moitié des résolutions considérées et à au moins une des étendues.

On remarque, au niveau de la zone 4 (Figure 7), que sur les 55 variables sélectionnées au départ, seules 21 contribuent de manière significative au moins une fois (règle 1).

	altitude	exposition (3*3)	exposition (5*5)	exposition (9*9)	penne (9*9)	rugosité du relief (3*3)	rugosité du relief (5*5)	rugosité du relief (20*20)	rugosité du relief (30*30)	relief local relatif (9*9)	relief local relatif (20*20)	relief local relatif (30*30)	durée d'insolation théorique	distance aux tourbières	distance aux crêtes	distance aux zones humides	type de formations superficielles	conditions d'utilisation collective	type de pâturage	utilisation	type de pratiques agricoles			
z4nI																						20 m		
																							30 m	
																							40 m	
																							50 m	
																							60 m	
																								70 m
																								80 m
z4nII																							90 m	
																							100 m	
																							20 m	
																							30 m	
																							40 m	
																							50 m	
																							60 m	
z4nIII																							70 m	
																							80 m	
																							90 m	
																							100 m	
																							20 m	
																							30 m	
																							40 m	
z6nIV																							50 m	
																							60 m	
																							70 m	
																							80 m	
																							90 m	
																							100 m	
																							100 m	

Figure 7. Sélection de variables dominantes en fonction de leur contribution aux différentes résolutions et étendues (zone 4)

Au final, 7 variables peuvent être qualifiées de « dominantes » (règle 2). L'application de cette méthode sur l'ensemble des zones, à toutes les échelles, permet de sélectionner 9 autres variables dominantes. Au totale, nous en avons donc 16 : altitude, exposition 5×5 et 9×9, rugosité du relief 30×30, relief local relatif 30×30, radiation solaire théorique, durée d'insolation théorique, distance aux tourbières, distance aux talwegs, distance aux crêtes, distance aux zones humides, formations superficielles, parcellaire, fréquentation animale, utilisation, exploitation des surfaces. Bien sûr, cette sélection mérite validation. Une base de données de grande taille a été construite sur l'ensemble du territoire à la résolution 20 m (les meilleurs résultats ayant été obtenus aux plus fortes résolutions). Pour l'alléger, nous avons écarté les indices estimés pour des noyaux de 3 et 5 cellules, sauf pour l'exposition 5×5 puisque qu'il s'agit d'une variable dominante. Les valeurs de reclassement de la rugosité et le relief local relatif ont été doublées, car elles présentaient, pour de larges fenêtres, trop de classes pour Pegase. La base de données, comportant 33 variables explicatives plus la variable cible, a été analysée dans trois conditions en activant : avec les 33 variables explicatives (HC1), avec les 16 variables dominantes seules (HC2) et les 17 variables secondaires seules (HC3).

On note que la néguentropie est donc pratiquement identique avec deux fois moins de variables (HC1 et HC2) (Tableau 2). Ces valeurs sont elles-mêmes proches de celles observées pour la zone z4nIV (60 %). Il faut rappeler que les variables exclues de ces deux analyses contribuent, même de façon infime, à l'entropie du système, expliquant ainsi des valeurs légèrement plus faibles. À l'inverse, pour HC3, la redondance est seulement de 35,5 %. L'ensemble des variables sélectionnées cerne alors le système de manière insatisfaisante. Il s'agit là d'une première preuve de la validité de notre sélection.

	HC1	HC2	HC3
Entropie initiale	2,122		
Entropie finale	0,882	0,925	1,368
Néguentropie	1,241	1,197	0,754
Redondance	58,4 %	56,4 %	35,5 %

Tableau 2. Résultats des trois tests de validation HC1, HC2 et HC3

L'étude des contributions des variables explicatives sélectionnées (Figure 8) nous permet également d'abonder dans ce sens. En effet, 7 variables ont une contribution supérieure à 5 % dans la configuration HC1.

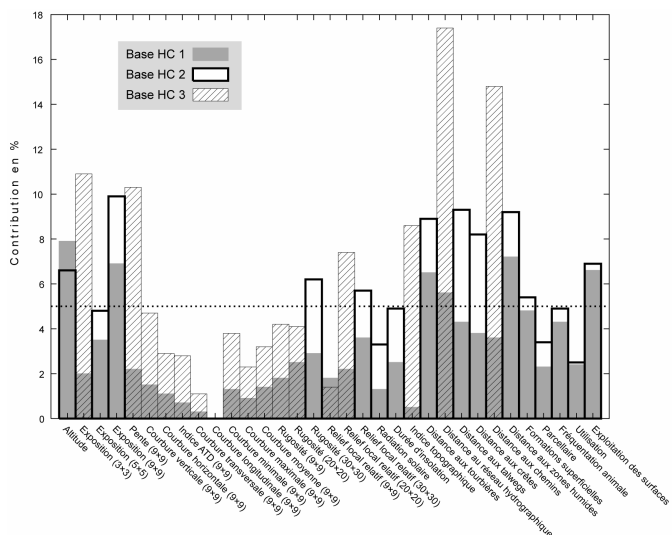


Figure 8. Contributions des différentes variables explicatives à la néguentropie pour les trois tests de validation HC1, HC2 et HC3

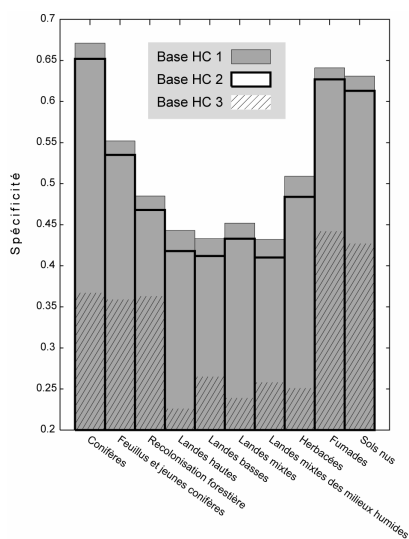


Figure 9. Spécificités des classes de la variables cibles pour les trois tests de validation HC1, HC2 et HC3

Toutes, sauf la « distance au réseau hydrographique » (proche des « distances aux talwegs »), ont été identifiées comme variables dominantes. Pour HC2, la contribution des variables dominantes augmente, sauf l'altitude. Nous mettons alors en évidence, lors de l'utilisation de toutes les variables, des phénomènes de masquage des variables dominantes par les variables secondaires. La comparaison des spécificités des classes de la variable cible montre des valeurs proches obtenues avec les configurations HC1 et HC2, mais une nette dégradation des résultats pour HC3 (Figure 9). Nous rajoutons ici un autre indicateur permettant de comparer les trois traitements. Il s'agit de l'étude des cartes de l'état prédit (Figure 10).

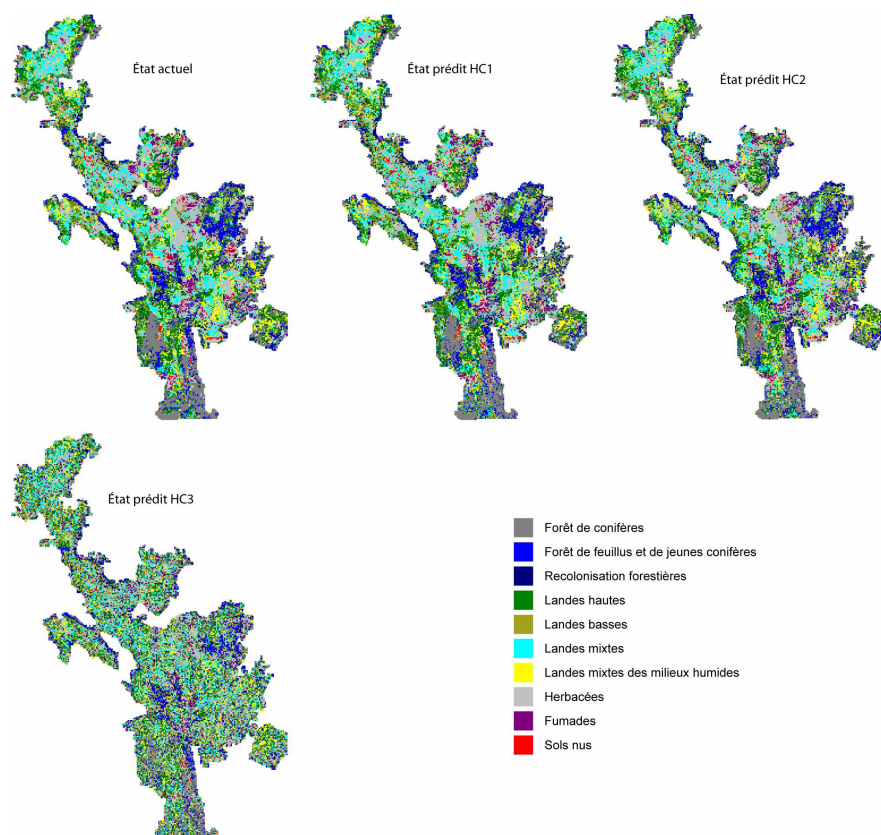


Figure 10. Cartographies des états prédits pour les trois tests de validation HC1, HC2 et HC3

Dans les cas HC1 et HC2, on observe des cartographies de l'état prédit très proche de l'état actuel, traduisant une bonne perception des liens existants entre combinaison de variables explicatives et classes de la variable cible, et donc une bonne modélisation. Par contre, on remarque clairement une nette dégradation avec la configuration HC3 se manifestant par une apparence « pointilliste ». Ainsi, que l'on analyse la néguentropie, les contributions, les spécificités ou bien la cartographie de l'état prédit, nous pouvons valider notre sélection de variables dominantes.

5. Discussion

5.1. Les résultats

L'importance d'un contexte multiéchelle lors de l'utilisation de Pegase, et plus généralement en analyse spatiale, est démontrée. Malgré l'homogénéité du territoire, chaque zone présente son organisation propre. Une simple étude de l'ensemble de ce territoire ne saurait permettre l'extrapolation à une zone plus restreinte. À l'inverse, l'étude d'une seule zone ne saurait expliquer l'organisation de la totalité du territoire. Notre méthodologie conduit à la mise en évidence de variables dominantes qui permettent de réduire la taille de la base de données et de s'affranchir des effets de masquage entre variables. Leur non identification au préalable peut conduire à sous-estimer leur(s) influence(s). Au mieux, l'organisation du paysage a été soulevée à 60 %. Bien qu'important, ce résultat reste insuffisant puisque 40 % ne sont pas expliqués. Il est certain que le hasard intervient dans l'organisation du système mais il ne saurait en expliquer à lui seul 40 %. Une autre explication possible est l'existence d'autres variables structurantes, non prises en compte au cours de notre analyse. En effet, le paysage se forme et change en fonction d'une grande diversité de facteurs inhérents à différents domaines (écologie, climat, histoire, économie...). L'idéal serait d'approcher cette globalité au plus près possible mais, dans cette tentative, on se heurterait à des obstacles multiples et difficiles à surmonter (comme cerner les influences du micro-climat). Il est possible également que la précision et la validité de certaines variables explicatives soient insuffisantes (par exemple celles liées aux pratiques pastorales). Il en va de même pour la variable cible, la cartographie de l'occupation du sol présentant des confusions qui influencent la procédure divisive (Mazagol 2006).

5.2. La méthodologie

Notre méthodologie présente certaines limites :

– la sélection des zones, de leurs étendues, de leurs résolutions peut prêter à discussion. Malgré cela, certaines constatations restent valables pour chacune d'elles, légitimant par la même occasion les résultats obtenus. Par exemple, le comportement des spécificités reste sensiblement identique d'une zone à l'autre. Cela tend donc à valider là encore nos résultats ;

– la zone de plus haut niveau d'étendue est commune aux zones de niveau I amenant à s'interroger sur l'impact de la variation de l'étendue sur les résultats. Notre choix induit que plus cette dernière augmente, plus on tend vers une seule et même zone. Ceci dit, elle correspond au seul endroit des Hautes Chaumes suffisamment haut et large ;

– nous n'avons pas testé notre méthodologie sur un autre terrain d'étude. Appliquer ces traitements sur un territoire alternatif, avec éventuellement des étendues plus importantes, une structure différente, soumis à des contraintes variées permettrait de confirmer nos résultats ;

– nous n'avons pas testé l'influence que pourraient avoir les variations de stratégie de reclassements. Par exemple, les courbures ont été reclassées de manière la plus simple qu'il soit, faisant apparaître les zones à courbures négatives, positives et nulles. De même, les pentes ou les distances par rapport aux crêtes ou aux talwegs ont été reclassées de manière linéaire. Cela mérite réflexion.

D'autre part, Pegase mériterait d'être intégré plus solidement encore au cœur de systèmes d'information géographique. Ainsi, nous envisageons le développement de l'outil en temps qu'extension ou module de logiciels SIG. La première étape serait une « fusion » des deux outils que nous avons développés : Asc2Peg et XLPeg, le second n'étant qu'une solution transitoire. L'existence d'un logiciel unique, permettant la structuration des bases de données Pegase, serait un gain de temps appréciable. D'autre part, la phase de préparation des données est longue et fastidieuse : rééchantillonnage, extraction des zones aux différentes étendues, exportation au format ascii... Cette étape pourrait, elle aussi, être automatisée. Il en va de même pour l'exportation des résultats qui pourrait faire l'objet d'une visualisation, directement en sortie, avec Gnuplot.

6. Conclusion

L'intérêt de l'algorithme Pegase est ici à nouveau démontré, d'autant plus qu'il est intégré au sein d'une méthodologie multiéchelle ancrée dans un contexte SIG. L'association se révèle donc efficace pour sélectionner des variables dominantes dans un paysage. Mais l'utilisation reste difficile en l'état actuel des choses. Ainsi, il est à présent nécessaire de l'intégrer de manière plus étroite au sein d'un SIG, par le développement d'une extension ou d'un module à un logiciel couramment utilisé.

Cela serait d'autant plus intéressant que notre méthode à montrer d'autres potentialités (Mazagol 2006, 2007). Elle permet par exemple de mieux cerner les

processus écologiques ayant cours au sein du milieu étudié, tel un phénomène de recolonisation forestière, ou bien l'existence d'un gradient de biomasse au sein des classes de la variable cible. Mais surtout, il est possible la placer comme outil potentiel d'aide à la mise en place de cartographies par télédétection, en utilisant la notion de spécificité ainsi que l'exploitation et l'analyse de confusions qui peuvent apparaître au cours de l'exécution de l'algorithme.

Allen T. F. H., Star T. B., 1982, "*Hierarchy: perspectives for ecological complexity*", The University of Chicago Press, Chicago.

Baudry J., 1992, « Approche spatiale des phénomènes écologiques : détection des effets d'échelles » in Auger P. et al. (eds.), *Hierarchies et échelles en écologie*, Naturalia Publications, Turriers, 157-171.

Baulig H., 1928, Le plateau central de la France et sa bordure méditerranéenne : Étude morphologique, Thèse de doctorat, Paris.

Bessenay C., 1995, La gestion des données environnementales dans un espace naturel sensible : le Système d'Information Géographique des Hautes-Chaumes foréziennes, Thèse de doctorat de 3^{ème} cycle, Université Jean Monnet, Saint-Etienne.

Casado M. A., Ramirez-Sanz L., de Miguel J. M., de Pablo C. L., 1997, An objective method for partitioning dendograms based on entropy parameters, *Plant Ecology*, 131, n° 2, 193-197.

Couhert J.-P., 1988, Les Hautes Chaumes des Monts du Forez : bilan de l'utilisation et des perspectives d'aménagement d'un territoire, Thèse de Doctorat de 3^{ème} cycle, Université de Clermont-Ferrand, Clermont-Ferrand.

Cubizolle H., 2005, Paléoenvironnement et turfigénèse, l'apport de l'étude des tourbières à la connaissance de l'évolution holocène des milieux montagnards du Massif Central oriental granitique, Habilitation à Diriger des Recherches, Université de Limoges, Limoges.

D'Alverny A., 1907, « Les HautesChaumes du Forez », *Bulletin de la Diana*, 11, 160-211.

Damon M., 1972, Les Jasseries des Monts du Forez. Sociologie de la vie pastorale, Thèse de Doctorat de 3^{ème} cycle, Université Lyon II, Lyon.

Davis F. W., Dozier J., 1990, « Information analysis of a spatial data base for ecological land classification », *Photogrammetric engeneering and remote sensing*, 56, n° 5, 605-613.

Delpoux M., 1972, « La science du paysage, une science diagonale », *Revue géographique des Pyrénées et du sud-ouest*, 4, n° 2, 127-134.

Dumanski J., Phipps M., Huffman E., 1987, "A study of the relationships between sol survey data and agricultural land use using information theory", *Canadian Journal of Sol Sciences*, 67, 95-102.

Ernoul A., Bureau F., Poudevigne I., 2003, "Patterns of organisation in changing landscapes: implications for the management of biodiversity", *Landscape Ecology*, 18, n° 3, 239-251.

Etlicher B., 1986, Les massifs du Forez, du Pilat et du Vivarais : régionalisation et dynamique des héritages glaciaires et périglaciaires en moyenne montagne cristalline, Thèse de Doctorat d'Etat, Université Lyon II, Lyon.

Etlicher B., Bessenay C., Couhert J.-P., Faury O., Francez A.-J., Suchel J.-B., Thébaud G., 1993, *Les Hautes Chaumes du Forez : Diagnostic écologique pour la gestion d'un espace sensible*, Publications de l'université de Saint-Étienne, Saint-Étienne.

Farina A., 1998, *Principles and methods in landscape ecology*, Chapman and Hall, Cambridge.

Fel A., 1962, Les Hautes Terres du Massif Central, tradition paysanne et économie agricole, Thèse de doctorat d'état, Université de Clermont-Ferrand, Clermont-Ferrand.

Jacqueminet C., Etlicher B., Sourp E., Laire R., 2002, « Un système d'information pour la concertation et l'aide à la décision sur un site Natura 2000 dans le Parc naturel régional Livradois Forez », *Géoévénement 2002*, Paris, 9-11 avril 2002.

Janssen C. R., 1990, « Histoire de la végétation des Monts du Forez », in Berger et al. (ed.) *Les Monts du Forez : le milieu et les hommes*, Centre d'Etudes Foréziennes, Saint-Étienne, 65-79.

Jelinski D., Wu J., 1996, "The modifiable areal unit problem and implications for landscape ecology", *Landscape Ecology*, 11, n° 3, 129-140.

Lundström-Gilliéron C., 1999, Modélisation de la dynamique du paysage : outil d'aide à la décision pour une gestion du territoire tenant compte de la faune. Étude de cas : le lièvre en Suisse., Thèse de doctorat, École Polytechnique Fédérale de Lausanne, Lausanne.

Lundström-Gilliéron C., Schlaepfer R., 2003, Hare abundance as an indicator for urbanisation and intensification of agriculture in Western Europe, *Ecological Modelling*, 168, n° 3, 283-301.

Mazagol P.-O., 2003, « Influence des variables physiographiques sur l'organisation d'un paysage », *14ème Rencontres Régionales de la Recherche*, Région Rhône-Alpes, Saint-Étienne, 21 Octobre 2003.

Mazagol P.-O., 2007, « Mise en évidence de « variables dominantes » dans l'organisation d'un paysage en utilisant la notion d'entropie spatiale : approche multiéchelles », *Actes des Cinquièmes Rencontres de ThéoQuant*, Université de Franche-Comté, Besançon, 10-12 janvier 2007.

Mazagol P.-O., 2006, Recherche de variables contribuant à l'organisation d'un paysage de moyenne montagne : le cas des Hautes Chaumes du Forez, Thèse de doctorat de 3ème cycle, Université Jean Monnet, Saint-Étienne.

Mazagol P.-O., 2005, XLPEG : Outil de mise en place d'entêtes de bases de données Pegase, CRENAM, Université Jean Monnet.

Mazagol P.-O., Mayoud R., 2005, Asc2Peg : outil de construction de bases de données Pegase à partir de fichiers au format d'échange ESRI ASCII Grid, CRENAM, Université Jean Monnet,

Mazagol P.-O., Porteret J., 2006a, « Étude de l'organisation du couvert végétal des Hautes Chaumes du Forez : un exemple d'utilisation de l'algorithme Pegase sur le secteur tourbeux de Gourgon-Bazanne pour l'année 1962 », *Colloque RTP "MoDyS"*, CNRS, Lyon, 8-9 novembre 2006.

Mazagol P.-O., Porteret J., 2006b, « Étude de l'organisation du couvert végétal des Hautes Chaumes du Forez : un exemple d'utilisation de l'algorithme Pegase sur le secteur tourbeux de Gourgon-Bazanne pour l'année 2001 », *Colloque RTP "Paysage et environnement"*, CNRS, Chilhac, 27-30 septembre 2006.

Phipps M., 1981, "Entropy and community pattern analysis", *Journal of theoretical biology*, 93, 253-273.

Phipps M., 2003, « Espace, entropie, environnement », *Actes des Cinquièmes Rencontres de ThéoQuant*, Université de Franche-Comté, Besaçon, 22-23 février 2001.

Phipps M., 2002, PEGASE operation manual, Université d'Ottawa.

Phipps M., Baudry J., Burel F., 1986a, « Dynamiques de l'organisation écologique d'un paysage rural : modalité de la désorganisation dans une zone péri-urbaine », *Compte rendus de l'Académie des Sciences de Paris*, 303, Série III, 263-268.

Phipps M., Baudry J., Burel F., 1986b, « Ordre topo-écologique dans un paysage rural : les niches paysagiques », *Compte rendus de l'Académie des Sciences de Paris*, 302, Série III, 691-696.

Porteret J., 2001, Le rôle des sociétés humaines dans la mise en place et le développement des tourbières du Haut-Forez, Mémoire de Maîtrise, Université Jean Monnet, Saint-Étienne.

Thébaud G., 1990, « La végétation du Haut-Forez », in Berger et al. (ed.) *Les Monts du Forez : le milieu et les hommes*, Université Jean Monnet, 37-63.

Thébaud G., 1988, Le Haut-Forez et ses milieux naturels. Apport de l'analyse phytosociologique pour la connaissance écologique et géographique d'une moyenne montagne cristalline subalpine, Thèse de Doctorat d'état, Université Clermont-Ferrand II, Clermont-Ferrand.

Turner G. M., Gardner H. R., 1990, "Quantitative methods in Landscape Ecology: an introduction", in Turner G. M., Gardner H. R. (eds.), *Quantitative methods in Landscape Ecology*, Springer, New York, 3-14.

Turner M. G., O'Neill R. V., Gardner R. H., Milne B. T., 1989, "Effects of changing spatial

scale on the analysis of landscape pattern", *Landscape Ecology*, 3, n° 3-4, 153-162.

Wu J., Jelinski D. E., Luck M., Tueller P. T., 2000, « »Multiscale analysis of landscape heterogeneity: scale variance and pattern metrics« », *Geographique information sciences*, 6, 6-19.

Wu J., Shen W., Sun W., Tueller P. T., 2002, "Empirical patterns of the effects of changing scale on landscape metrics", *Landscape Ecology*, 17, n° 8, 761-782.