HAL
open science

Structures de haies dans un paysage agricole : une étude par chemin de Hilbert adaptatif et chaînes de Markov

Sébastien da Silva, Florence Le Ber, Claire Lavigne

To cite this version:

Sébastien da Silva, Florence Le Ber, Claire Lavigne. Structures de haies dans un paysage agricole : une étude par chemin de Hilbert adaptatif et chaînes de Markov. EGC 2016 – 16èemes Journées Francophones "Extraction et Gestion des Connaissances", Jan 2016, Reims, France. pp.279–290. hal-01266344

HAL Id: hal-01266344

https://hal.science/hal-01266344v1

Submitted on 5 Nov 2019

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.

Structures de haies dans un paysage agricole : une étude par chemin de Hilbert adaptatif et chaînes de Markov

Sébastien da Silva*, Florence Le Ber**
Claire Lavigne***

*LORIA et INRIA-NGE, Nancy
sebastien.da-silva@loria.fr,
**ICube, Université de Strasbourg, ENGEES, CNRS, Strasbourg
florence.leber@icube.unistra.fr
***INRA UR 1115, Avignon
claire.lavigne@avignon.inra.fr

Résumé. Dans cet article nous présentons une approche couplant une courbe remplissant l'espace et une chaîne de Markov pour analyser des données spatiales concernant la localisation de haies. Du fait de l'hétérogénéité spatiale des données, nous utilisons une courbe adaptative de Hilbert qui permet de linéariser l'espace en s'ajustant localement à la densité des données. Pour ensuite exploiter la séquence produite, il est nécessaire de caractériser la distance entre un point et son prédecesseur sur la courbe ainsi que la densité locale. Nous proposons de calculer un temps d'accès à un point à partir du point précédent en utilisant la notion de profondeur de découpe. Cette variable, couplée avec les variables caractérisant les haies est ensuite analysée avec un modèle de Markov. Nous présentons et interprétons les résultats obtenus sur un jeu de données d'environ 10000 segments de haies d'une zone de la Basse vallée de la Durance.

1 Introduction

La structure spatiale des paysages, c'est-à-dire leur composition et leur configuration, affecte de nombreux processus qui se déroulent à cette échelle, qu'ils soient par exemple écologiques, épidémiques, hydrologiques, ou culturels... Le champ de la modélisation des paysages s'est donc considérablement développé récemment, en particulier pour concevoir des paysages multifonctionnels. Si beaucoup d'études se sont concentrées sur la modélisation des éléments surfaciques (parcelles agricoles, forêts...) moins d'études ont porté sur les éléments linéaires tels que les haies. Cependant, ces éléments sont primordiaux pour la conservation de certaines espèces qui vivent dans ces paysages en jouant un rôle d'habitat, de corridor ou au contraire de barrière au déplacement (Davies et Pullin, 2007). Les haies sont également importantes pour lutter contre l'érosion des sols (Lacoste et al., 2015) et limiter la pollution des eaux de surface (Borin et al., 2010). Il est donc important, comme cela a été fait pour les éléments surfaciques (Castellazzi et al., 2007; Lazrak et al., 2010), de chercher à *(i)* caractériser les structures spatiales des haies, *(ii)* relier ces structures à des fonctions vis-à-vis des processus

bio-physiques en cours dans le paysage, *(iii)* développer des méthodes de simulation afin d'explorer différentes configurations permettant de favoriser ou non telle ou telle fonction. A notre connaissance, il n'existe pas de travaux permettant de répondre à ces différents points et particulièrement au point *(iii)*. Nous avons donc développé une méthode de caractérisation des structures spatiales de haies, dans l'objectif de construire un modèle permettant de générer des structures aléatoires utilisées ensuite en entrée de modèles de processus bio-physiques. Dans cet article, nous tentons de montrer que cette méthode arrive bien à distinguer et caractériser différentes structures de haies que nous pouvons rattacher à des fonctions agricoles.

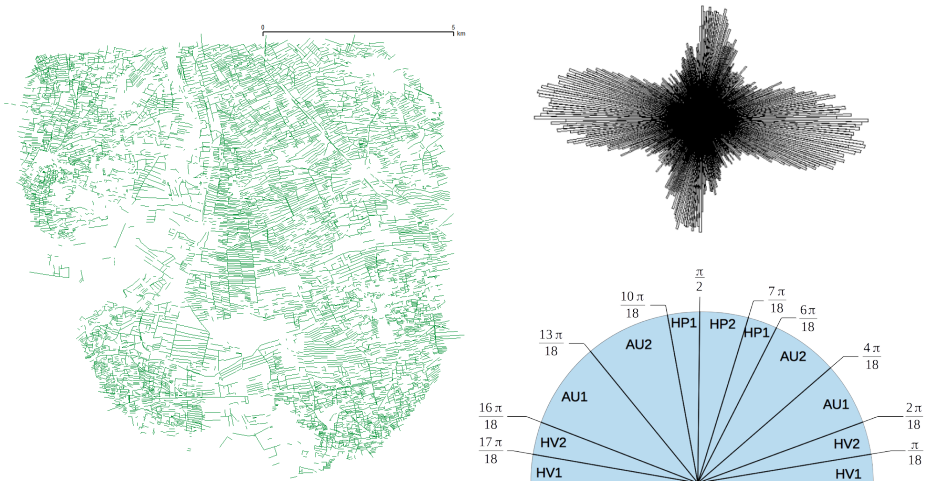
Pour caractériser ces structures, et plus largement celles des linéaires dans les paysages, nous adoptons une démarche de fouille de données spatiales hétérogènes, en nous inspirant des travaux menés en traitement d'images (Benmiloud et Pieczynski, 1995) et plus récemment pour l'étude d'organisations spatiales agricoles (Lazrak et al., 2010). Les caractéristiques qui nous intéressent principalement sont des caractéristiques propres à chaque haie (longueur, orientation) et des caractéristiques inter-haies, par exemple la relation entre les orientations d'un groupe de haies voisines. Du fait de l'hétérogénéïté spatiale des données, la méthode développée procède d'abord en une linéarisation de l'espace grâce à une courbe remplissant l'espace. L'utilisation d'une telle courbe a deux objectifs principaux : *(i)* ordonner de manière unidimensionnelle les barycentres des segments de haies présents dans une portion d'espace (appelée cellule), et donc transformer une information 2D en une information 1D ; *(ii)* respecter au mieux la notion de proximité spatiale entre les barycentres (propriété de localité). L'ordonnancement uni-dimensionnel ainsi réalisé permet d'utiliser les méthodes d'apprentissage de type Markov dont les algorithmes sont bien connus et qui, de plus, sont capables générer facilement des données aléatoires après apprentissage.

Le plan de l'article est le suivant. Dans la section 2 sont présentés le jeu de données sur lequel nous avons testé notre approche ainsi que les prétraitements réalisés. La section 3 décrit brièvement les méthodes employées. Dans la section 4, Les résultats obtenus sont présentés et discutés. Notre travail est situé par rapport à l'existant dans la section 5 puis nous concluons en proposant quelques pistes dans la perspective de générer aléatoirement des structures de haies comme supports pour la simulation de processus bio-physiques.

2 Présentation des données

Notre travail porte sur un jeu de données représentant des linéaires de haies dans le Sud-est de la France (zone d'étude Basse vallée de la Durance, coordonnées en WGS84 : de 43°46'27" N à 43°51'23" N et de 4°51'12" E à 4°57'34" E). Cette région est caractérisée par une forte densité de haies brise-vent de cyprès ou peupliers orientées Est-Ouest pour protéger les cultures du mistral. Les autres haies présentes dans le paysage sont pour la plupart des haies de platanes en bord de route ou de fossés ou des haies composites en bord de parcelles. Le linéaire de haies a été numérisé à la main à partir de photos aériennes. Les routes et canaux ont également été numérisés mais cette information ne sera pas traitée ici. Le jeu de données est donc constitué de 11557 segments linéaires de haies (voir figure 1(a)) caractérisés chacun par la position de leur centre, leur longueur et leur orientation.

Cette caractérisation met bien en évidence la présence d'une forte proportion de haies brise-vents (HV), ainsi qu'un grand nombre de haies qui leur sont perpendiculaires (HP) et une classe intermédiaire (AU) (voir (Da Silva et al., 2014) pour plus de détails). Ceci nous



(a) Données (b) Orientations

FIG. 1 – *Données d'origine, seules les informations sur les haies sont représentées (a), distribution des valeurs d'orientation et découpage en classes (b)*

a conduit à définir six classes d'orientation par rapport à l'axe Nord-Sud (modulo π, voir la figure 1(b) en bas) :

- (HV1) $[0, \pi/18[\cup[17\pi/18, \pi[$; (HV2) $[\pi/18, 2\pi/18[\cup[16\pi/18, 17\pi/18[$;
- (AU1) $[2\pi/18, 4\pi/18[\cup[13\pi/18, 16\pi/18[$; (AU2) $[4\pi/18, 6\pi/18[\cup[10\pi/18, 13\pi/18[$;
- (HP1) $[6\pi/18, 7\pi/18[\cup[\pi/2, 10\pi/18[$; (HP2) $[7\pi/18, \pi/2[$.

Ces classes comprennent respectivement 53,8% (HV1 et 2 confondues), 18,4% (HP1 et 2 confondues), 27,8% (AU1 et 2) des segments (voir figure 1(a) à droite). La longueur moyenne des haies est de 101,5 m $\pm$ 79,8. Bien que la distribution des longueurs des segments ne fasse pas apparaitre de classes aussi clairement que pour leurs orientations, celles-ci ont été regroupées en 6 classes, comportant chacune 16,7% des segments (bornes : 36,1 m ; 57,6 ; 82,5 ; 112,3 ; 160,9). Il n'existe pas de corrélation forte entre la longueur des segments et leur orientation (r= -0,11).

La distribution spatiale des différents types de segments est hétérogène sur le paysage, certaines zones présentant visuellement de plus fortes densités de haies que d'autres. Pour pouvoir travailler sur différents types de distributions spatiales de haies – et en même temps réduire la taille des données, la zone d'étude a donc été découpée en 121 cellules carrées de 1100 m de côté (voir figure 2 à gauche). Le choix de la maille a été dicté par la nécessité d'avoir un nombre suffisant de segments par cellule (de l'ordre de la centaine) tout en ayant une homogénéité de la distribution spatiale des haies à l'intérieur de chaque cellule et une hétérogénéité entre cellules. Parmi celles-ci, 64 cellules, situées en totalité à l'intérieur de la zone numérisée, ont été conservées pour la suite des analyses.

Nous nous intéressons particulièrement à la distribution spatiale des différents types de haies dans ce paysage. Pour l'appréhender, nous avons opéré une classification ascendante hiérarchique (fonction hclust, package Stats, logiciel R Commander V1.8-1) des cellules selon

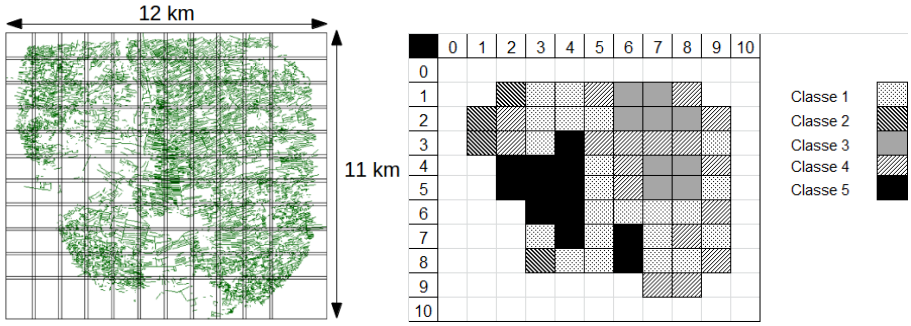


FIG. 2 – *Découpage en cellules et classification des cellules – la numérotation des cellules s'effectue en accollant numéro de ligne puis numéro de colonne*

leurs nombres de haies HP, HV et AU. Cette classification a abouti à 5 classes de cellules comportant respectivement 22, 4, 9, 17 et 12 cellules et en moyenne ($\pm$ écart-type) 107$\pm$15, 265$\pm$18, 163$\pm$16, 165$\pm$15 et 51$\pm$10 segments de haies. Les classes 1, 3 et 4 sont celles qui contiennent proportionnellement le plus de haies HV (environ 60%), elles correspondent grossièrement à des zones de vergers plus ou moins denses (mixés avec des cultures pour la classe 1). La classe 2 contient proportionnellement le plus de haies HP (30%), la classe 3 le moins (5%). La classe 5 correspond à une zone mixte, comprenant à l'ouest de la garrigue, des vignes et des oliviers tandis qu'au sud se trouvent des cultures de plein champ, où les haies sont peu présentes.

3 Méthodes

Pour exploiter les données, nous utilisons une technique de linéarisation qui permet de transformer une information 2D en 1D et d'utiliser ensuite des algorithmes classiques d'apprentissage comme les chaînes de Markov. Les courbes remplissant l'espace sont utilisées pour linéariser une information spatiale et sont couramment employées en traitement d'images mais aussi pour l'étude de phénomènes spatiaux. On utilise en particulier la courbe de Hilbert-Peano qui s'appuie sur un motif simple et permet de préserver une partie de l'information spatiale (les points voisins sur la courbe sont voisins dans l'espace). Cependant, si le phénomène spatial étudié présente des irrégularités, des zones très denses et des zones vides d'information, la courbe classique, régulière, est peu appropriée. Il faut alors utiliser une méthode qui s'adapte localement à l'information, comme la courbe adaptative de Hilbert (CHA), dont l'algorithme a été décrit par Quinqueton et Berthod (1981).

3.1 Construction d'un chemin de Hilbert adaptatif

La méthode de construction d'un CHA prend en entrée une cellule carrée contenant un ensemble de points spatialement distribués, qui sont parcourus de manière déterministe. Sur la figure 3, est représentée une cellule comportant un ensemble de barycentres de segments de

haies (étape 0). Le point de départ de l'algorithme est cette cellule englobant tous les barycentres ainsi qu'une direction de parcours (ici de bas à gauche vers le haut puis bas à droite). La cellule comportant plus de deux points est découpée en quatre cases (étape 1) et on se place dans la case en bas à gauche. A une étape suivante de l'algorithme, on considère le nombre de barycentres dans la case courante : si ce nombre est égal à 1 ou 0, on passe à la case suivante ; sinon on découpe la case en quatre sous-cases et on se place dans la case en bas à gauche.

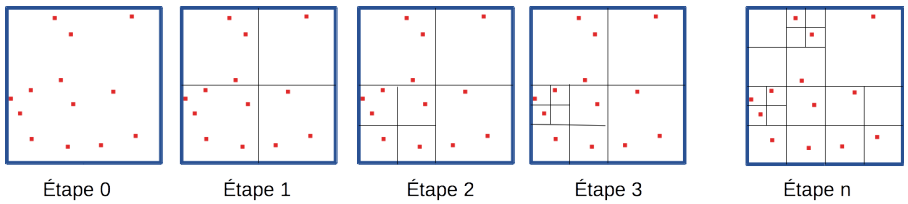


FIG. 3 – *Découpe d'une cellule selon l'algorithme du chemin de Hilbert Adaptatitif (CHA)*

Selon la dimension e de la cellule initiale, le nombre N de points et leur répartition dans cette cellule, la construction du chemin conduit récursivement à la définition de cases de plus en plus petites, jusqu'à une profondeur maximale $k_{\max}$. La complexité de l'algorithme est de l'ordre de $2^e N k_{\max}$ (Quinqueton et Berthod, 1981).

En sortie du processus, la cellule originelle est donc décomposée en cases, qui sont ensuite parcourues pour créer une séquence de points, ici de segments de haies caractérisés par une longueur et une orientation. Il est important dans notre étude de conserver l'information sur la distance entre les haies, c'est pourquoi nous cherchons à caractériser la structure du chemin construit.

3.2 Grandeurs caractéristiques et apprentissage

Différentes grandeurs peuvent être utilisées pour caractériser un CHA, en premier lieu les grandeurs associées aux objets d'intérêt (ici les segments de haie), et en deuxième lieu les grandeurs propres au chemin, principalement la taille des cases créées autour des différents points. Soit A la surface de la cellule originelle et A_i la surface de la case courante C_i ($i \in [1, N_C]$, N_C le nombre de cases). On définit alors la profondeur de découpe de la case C_i, notée PdD_i, par :

$$PdD_i = \frac{\ln(A) - \ln(A_i)}{2\ln(2)}$$

puis le temps de parcours de la case C_i, noté TdP_i, par :

$$TdP_i = \frac{1}{2^{PdD_i}}$$

et enfin le temps d'accès au point $k \in]1, N]$ de la case C_i depuis le point $k-1$ situé dans la case C_{i-n} ($n < i$) par :

$$TdA_k = \frac{TdP_{i-n}}{2} + \sum_{j=i-n+1}^{j=i-1} TdP_j + \frac{TdP_i}{2}$$

On considère que les cases de départ et d'arrivée sont parcourues à moitié et les cases intermédiaires C_j entièrement (cf. figure 4).

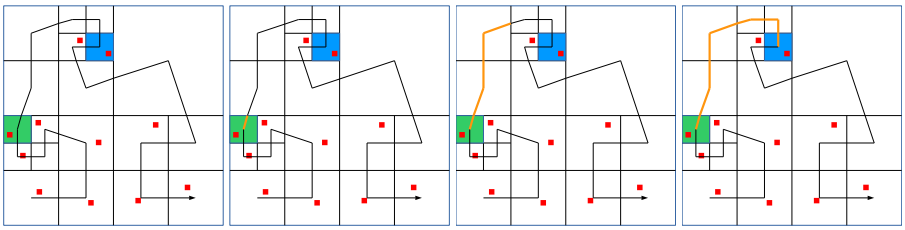


FIG. 4 – *Visualisation du temps d'accès au point $k = 7$ (case bleue) à partir du point $k - 1$ (case verte)*

Notons que pour un chemin de Hilbert classique (où plusieurs points peuvent se trouver dans une même case) toutes les cases ont la même surface et les différentes variables définies ci-dessus sont donc des constantes.

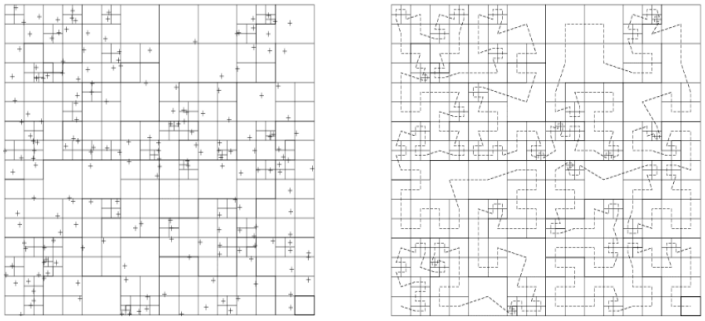
Finalement le chemin obtenu sur une cellule peut être transformé en une séquence de vecteurs $\vec{m}_k$ contenant les informations suivantes : temps d'accès depuis le point précédent (TdA_k), orientation du segment de haie (θ_k), longueur du segment de haie (L_k). Les différentes valeurs sont discrétisées (cf. section 2). Ce type de séquences peut être caractérisé grâce à des chaînes de Markov (Baum et Petrie, 1966). Nous avons ici effectué l'apprentissage de chacune des variables séparément, du fait de la faible corrélation des deux variables longueur et orientation.

4 Résultats et interprétation

4.1 Chemins

La figure 5(a) présente un exemple de découpe d'une cellule de notre jeu de données. Nous pouvons remarquer la disparité des tailles des cases créées, confirmant l'adaptation de la division à la densité du phénomène observé. Le parcours de ce découpage est présenté sur la figure 5(b). Le point de départ du chemin est situé dans la case en bas à gauche et le point d'arrivée dans la case en bas à droite.

Pour caractériser les chemins et les relier aux classes de cellules définies *a priori*, on calcule pour chaque chemin le nombre de cases aux différentes profondeurs. La distribution des moyennes obtenues pour les classes est représentée sur la figure 6 (à gauche). Les distributions sont globalement similaires pour les classes 1 à 4, les case de profondeurs 4 et 5 étant majoritaires. La classe 2 contient relativement plus de petites cases (profondeur 6), tandis que la classe 1 contient plus de grandes cases (profondeur 3). La classe 5 a un profil particulier, avec un nombre moyen élevé de cases de profondeur 3, supérieur au nombre moyen de cases de profondeur 5. Sur les exemples pris ici, la cellule de la classe 4 présente une densité élevée et une certaine régularité des structures (les cases se répartissent principalement en deux niveaux de découpe) tandis que la cellule de la classe 1 a une densité moindre, et présente une plus grande hétérogénéité des strucures (les cases se répartissent dans trois niveaux de découpe).



(a) Découpe de la cellule (b) Parcours des cases

FIG. 5 – *CHA dans une cellule où chaque segment de haie est représenté par son barycentre*

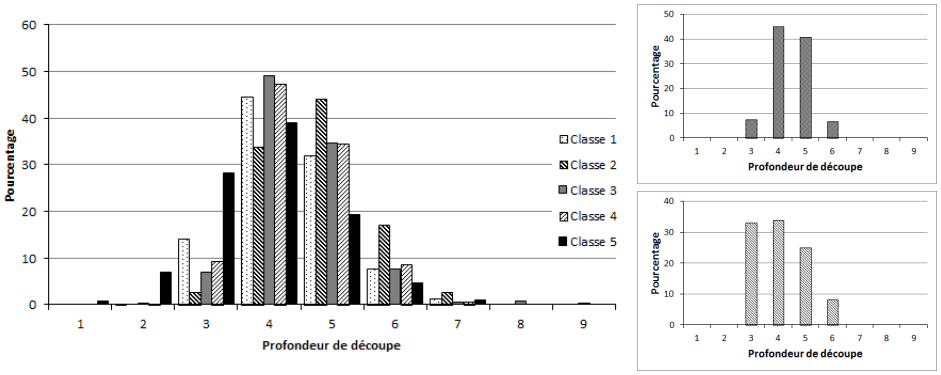


FIG. 6 – *Distribution du nombre moyen de cases aux différentes profondeurs pour les classes de cellules et exemple de deux cellules, en haut pour la classe 4, en bas pour la classe 1*

4.2 Chaînes de Markov

Pour chaque cellule, nous avons obtenu trois chaînes de Markov (sous la forme d'une liste d'états et d'une matrice de transitions entre états), une pour les orientations, une pour les longueurs, une pour les distances. Le nombre d'états est fixé à 6 pour les orientations et les longueurs (cf. section 2), alors qu'il est variable pour les distances.

Nous présentons ici sous forme de graphes (figure 7) les transitions entre états d'orientation obtenues pour les deux cellules exemples de la figure 6. Pour rappel, les 6 classes d'orientation correspondent à trois différents types de haies (d'un point de vue fonctionnel), haies brise-vents (HV1 et 2), haies de bord de route ou canal d'irrigation (HP1 et 2), perpendiculaires aux haies brise-vents, et haies autres (AU1 et 2).

Pour la cellule de la classe 1, on observe une prédominance des états HV1 et 2 (à relier au taux important de haies brise-vents), suivi des états AU puis des états HP. Les transitions de HP vers HV sont très importantes (60 et 80%) ce qui signifie que la majorité des haies de bord de route se trouve à proximité d'une haie brise-vents. Les haies HV sont voisines entre elles

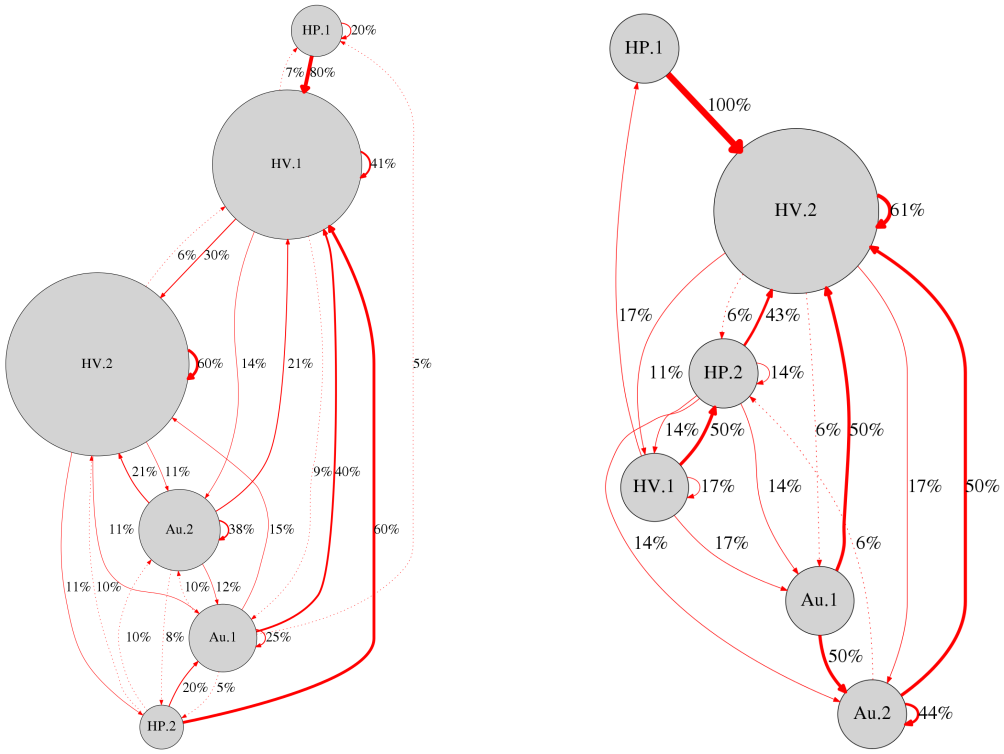


FIG. 7 – *Représentation des transitions entre états d'orientations obtenues pour une cellule de la classe 4 (à gauche) et une cellule de la classe 1 (à droite) ; les rayons des cercles sont proportionnels à la fréquence des états, l'épaisseur des lignes aux taux de transition*

(transitions de 30 à 60%), montrant une régularité de la disposition de ces haies même si leur orientation peut légèrement varier.

Pour la cellule de la classe 4, on observe une prédominance d'un seul état HV2 qui est fréquemment voisin de lui-même (transition de 61%) ; la relation entre l'état HP1 et cet état est très forte (transition de 100%), moins forte pour l'état HP2 (43%). Les haies brise-vents de la catégorie HV1 (orientation proche de l'axe Est-Ouest), peu nombreuses, avoisinent préférentiellement les haies HP2 (50%). Enfin les haies de type AU sont ici souvent à proximité d'elles-mêmes ou des haies HV2 (transitions de 50%).

Ces observations sont à relier aux structures différentes des deux cellules (cf. figure 8). La cellule de la classe 4 se trouvent dans une zone de vergers, les parcelles sont toutes bordées de haies brise-vents dont l'orientation varie un peu selon les réseaux de routes ou canaux. Cette zone est parcourue de quelques routes et canaux le long desquels on trouve des haies, sur lesquelles s'appuient les haies brise-vents. La cellule de la classe 1 est elle à cheval sur deux types de paysages, un paysage de vergers (proche de la cellule précédente) et un paysage de cultures, où les haies sont moins nombreuses et plus irrégulières dans leur orientation. Des

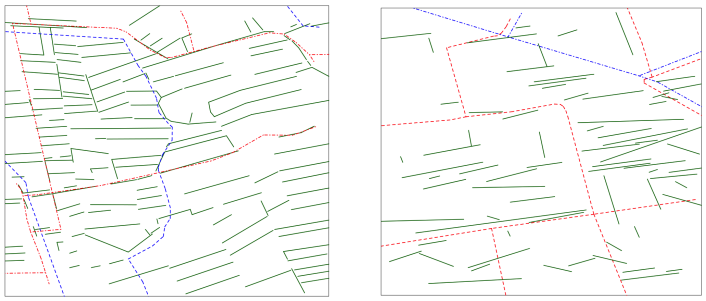


FIG. 8 – *Carte d'une cellule de la classe 4 (à gauche) et d'une cellule de la classe 1 (à droite) ; sont représentés les haies (en vert), canaux (en pointillés bleus) et routes (en pointillés rouges)*

résultats similaires sont trouvés sur les autres cellules et classes de ce jeu de données.

4.3 Discussion

L'étude a été menée sur un ensemble de 10.000 haies, que nous avons réparties en une centaine de cellules, dont 64 ont été étudiées. La classification *a priori* des cellules (sur le critère du nombre de haies) a permis de mettre en évidence des paysages types (zones de vergers plus ou moins denses, zones de cultures, zones de transition, zones d'habitat ...). L'analyse des cellules par l'application du chemin de Hilbert adaptatif et des chaînes de Markov permet d'affiner cette première classification en caractérisant les cellules par des critères structurels (distribution des cases dans les chemins, structures des matrices de transition) comme nous l'avons montré ci-dessus. Pour généraliser cette analyse, nous avons effectué une classification des matrices de transitions, en utilisant une distance entre matrices. La classification obtenue pour les matrices de transitions entre états d'orientations est présentée sur la figure 9.

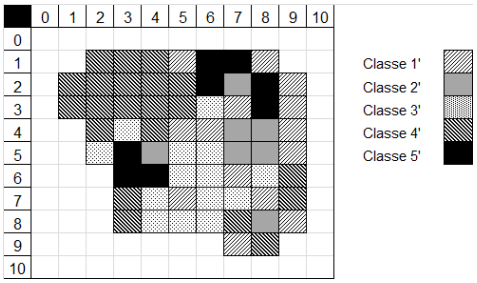


FIG. 9 – *Classification des cellules selon les matrices de transitions entre états d'orientations*

Nous avons choisi de constituer cinq classes afin de comparer cette classification à la classification originelle (sur le nombre de haies) présentée en figure 2. De manière globale on observe que ces classifications sont différentes, et donc que l'information apportée par la matrice (transitions entre les différents types de haies) est complémentaire de l'information sur

les seuls types de haie en présence dans une cellule. Plus particulièrement, considérons différentes sous-zones. Par exemple les quatre cellules 47, 48, 57, 58 (classes 3 et 2') sont stables d'une classification à l'autre : elles sont homogènes tant pour la densité des différents types de haies (HV prédominantes) que pour leur structure spatiale ; plus au nord en revanche, une route transversale vient perturber localement la structure (classe 3 puis classe 5'). Au nord-ouest, des cellules séparées dans la classification d'origine se trouvent regroupées (classe 4') : les densités sont variables mais la structure – de HV appuyées sur des HP – est identique. Il s'agit en effet d'une zone de maraîchage, bien distincte des zones de vergers plus à l'est.

5 Travaux connexes

Les modèles de Markov sont couramment utilisés pour la segmentation de données linéaires dans de nombreuses applications ; dans le domaine agro-environnemental qui nous intéresse, on peut citer par exemple la segmentation de données hydromorphologiques (Kehagias, 2004), la caractérisation de la distribution spatiale des couvertures du sol (Zhang et Li, 2005) ou encore la détermination de séquences d'occupations du sol (Le Ber et al., 2006). Les modèles de Markov sont également très utilisés en analyse d'image en association à des courbes remplissant l'espace qui permettent de transformer l'information 2D en 1D. Ces méthodes couplées sont en effet plus rapides que celles utilisant des modélisations par champs de Markov, tout en produisant de bons résultats (Benmiloud et Pieczynski, 1995). Les parcours d'images avec des courbes remplissant l'espace sont aussi utilisés pour la compression d'images (Kamata et al., 1993). Des courbes dites « adaptatées au contexte » ont été proposées par Dafner et al. (2000) avec pour objectif de parcourir une image en minimisant les passages d'un objet à l'autre. La courbe fractale est construite de bas en haut, à partir de petits parcours locaux qui sont ensuite fusionnés selon un ordre défini par la différence de valeurs entre les pixels de chaque parcours. Finalement la courbe obtenue permet de parcourir entièrement l'image en respectant au mieux la structure des objets qui la composent.

Concernant les données agricoles on peut citer les travaux de Lazrak et al. (2010) qui utilisent des modèles de Markov cachés, associés à des courbes de Hilbert-Peano pour analyser des cartes d'occupation du sol. La carte est parcourue de façon régulière et c'est la structure de la courbe qui détermine les points où on relève l'occupation du sol. Pour faire face à l'hétérogénéité des tailles de parcelles, ces auteurs ont fait le choix d'ajuster le niveau de découpe de la courbe afin de minimiser l'information délaissée. De plus, une portion de courbe située entièrement dans une parcelle (même occupation du sol) peut être réduite à un point – ce qui revient à construire une courbe partiellement adaptatée. Guyet (2010) exploite également des cartes d'occupations du sol, mais en utilisant les graphes reliant les barycentres des parcelles, graphes sur lesquels il recherche ensuite des motifs (occupations du sol voisines) fréquents.

Ce type de données spatiales est aussi exploité avec des approches statistiques ou géométriques. Les données présentées ici ont par exemple été étudiées avec des méthodes géométriques d'analyse de voisinage pour mettre en évidence la structuration de différents types de linéaires (Da Silva et al., 2014). Une telle approche ne permet pas toutefois d'obtenir un modèle du paysage comme nous le proposons ici.

Nous avons considéré les haies comme des segments autonomes – ce qui convenait à ce jeu de données spécifique. Dans d'autres paysages, les haies peuvent aussi être considérées dans leur ensemble, comme un réseau. De nombreuses méthodes existent pour caractériser

les réseaux – en particulier pour les réseaux urbains, largement étudiés en géographie. Des approches spécifiques ont été développées pour les réseaux naturels présents dans les paysages (voir par exemple (Thommeret et al., 2014) pour les réseaux hydrographiques).

6 Conclusion

L'approche développée ici a été appliquée sur deux jeux de données différenciés. Nous n'avons présenté qu'un jeu de données pour de questions de place et pour pouvoir développer l'interprétation des résultats. Les modèles construits pourront ensuite être utilisés dans la perspective de générer des structures de haies aléatoires – comme cela a été fait pour les parcellaires (Le Ber et al., 2009; Gaucherel et al., 2006) – qui puissent servir d'entrée à des modèles de processus bio-physiques. Ainsi les agronomes peuvent tester différentes configurations de paysages afin de répondre à des problématiques, par exemple, de protection des cultures, d'espèces ou de prévention de l'érosion des sols. Ce type de recherche a aussi pour objectif de concevoir des paysages agricoles mieux adaptés aux problématiques actuelles de l'agroécologie.

D'un point de vue méthodologique, l'approche développée s'est avérée pertinente pour mettre en évidence différents types de structures. Néanmoins il faut tester la stabilité des chemins construits et des chaînes de Markov qui en dérivent. En particulier, nous prévoyons de réaliser plusieurs parcours sur une même cellule afin d'obtenir des caractéristiques plus fiables.

Enfin l'approche que nous avons développée peut s'appliquer sur d'autres types de structures linéaires. En géographie, en particulier, on peut considérer des structures urbaines (bâtiments), rurales (terrasses, canaux) ou naturelles (reliefs érosifs ...).

Références

Baum, L. E. et T. Petrie (1966). Statistical Inference for Probabilistic Functions of Finite State Markov Chains. *The Annals of Mathematical Statistics 37*(6), 1554–1563.

Benmiloud, B. et W. Pieczynski (1995). Estimation des paramètres dans les chaînes de Markov cachées et segmentation d'images. *Traitement du signal 12*(5), 433–454.

Borin, M., M. Passoni, M. Thiene, et T. Tempesta (2010). Multiple functions of buffer strips in farming areas. *European Journal of Agronomy 32*(1), 103 – 111. Cropping Systems Design : new methods for new challenges.

Castellazzi, M., J. Perry, N. Colbach, H. Monod, K. Adamczyk, V. Viaud, et K. Conrad (2007). New measures and tests of temporal and spatial pattern of crops in agricultural landscapes. *Agriculture, Ecosystems & Environment 118*, 339–349.

Da Silva, S., F. Le Ber, et C. Lavigne (2014). Structure Analysis of Hedgerows with Respect to Perennial Landscape Lines in Two Contrasting French Agricultural Landscapes. *International Journal of Agricultural and Environmental Information Systems (IJAEIS) 5*(1), 19.

Dafner, R., D. Cohen-Or, et Y. Matias (2000). Context-based space filling curves. *Computer Graphics Forum 19*(3), 209–218.

Davies, Z. G. et A. S. Pullin (2007). Are hedgerows effective corridors between fragments of woodland habitat ? an evidence-based approach. *Landscape Ecology 22*(3), 333–351.

Gaucherel, C., D. Fleury, D. Auclair, et P. Dreyfus (2006). Neutral models for patchy landscapes. *Ecological Modelling 197*(1-2), 159–170.

Guyet, T. (2010). Fouille de données spatiales pour la caractérisation spatiale de paysages en lien avec des fonctionnalités agro-écologiques. In *Spatial Analysis and GEOmatics (SAGEO'10)*, pp. 3.

Kamata, S., R. Eason, et E. Kawaguchi (1993). An implementation of the Hilbert scanning algorithm and its application to data compression. *IEICE Transaction information and systems E-76*(4), 420–427.

Kehagias, A. (2004). A hidden Markov model segmentation procedure for hydrological and environmental time series. *Stochastic Environmental Reseau 18*, 117 – 130.

Lacoste, M., V. Viaud, D. Michot, et C. Walter (2015). Landscape-scale modelling of erosion processes and soil carbon dynamics under land-use and climate change in agroecosystems. *European Journal of Soil Science 66*(4), 780–791.

Lazrak, E., J.-F. Mari, et M. Benoît (2010). Landscape regularity modelling for environmental challenges in agriculture. *Landscape Ecology 25*, 169–183.

Le Ber, F., M. Benoît, C. Schott, J.-F. Mari, et C. Mignolet (2006). Studying crop sequences with CARROTAGE, a HMM-based data mining software. *Ecological Modelling 191*(1), 170–185.

Le Ber, F., C. Lavigne, K. Adamczyk, F. Angevin, N. Colbach, J.-F. Mari, et H. Monod (2009). Neutral modelling of agricultural landscapes by tessellation methods – application for gene flow simulation. *Ecological Modelling 220*, 3536–3545.

Quinqueton, J. et M. Berthod (1981). A Locally Adaptive Peano Scanning Algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-3*(4).

Thommeret, N., J.-S. Bailly, J.-M. Bardet, B. Kaiser, et C. Puech (2014). Dimensions fractales de réseaux vectoriels : méthodes d'estimation et robustesse des résultats. *Cybergeo : European Journal of Geography [Online], Systems, Modelling, Geostatistics*. Document 688, en ligne depuis le 18 août 2014.

Zhang, C. et W. Li (2005). Markov Chain Modeling of Multinomial Land-Cover Classes. *GIScience & Remote Sensing 42*(1), 1–18.

Summary

This article presents an approach coupling space filling curves and Markov chains for analysing spatial data about hedge localisation. Due to the data spatial heterogeneity, an adaptative Hilbert curve is used, allowing space linearisation with respect to data local density. In order to exploit the resulting sequence, we need to characterise both the distance between a point and its predecessor on the curve and the local density. We propose therefore to compute an access time for a point from its predecessor based on the cutting depth notion. This variable, together with the hedge characterising variables are then analysed through a Markov model. Results obtained on a dataset of about 10000 hedge segments from a zone in the Durance low valley are presented and discussed.