

# Fouille de données à l'aide de modèles stochastiques : segmentation temporo-spatiale des successions de cultures d'un territoire agricole à l'aide de HMM2

El-Ghali Lazrak, Marc Benoît, Jean-Francois Mari

## ► To cite this version:

El-Ghali Lazrak, Marc Benoît, Jean-Francois Mari. Fouille de données à l'aide de modèles stochastiques : segmentation temporo-spatiale des successions de cultures d'un territoire agricole à l'aide de HMM2. STIC 2009 Environnement, Jun 2009, Calais, France. 2009. <inria-00398182>

**HAL Id: inria-00398182**

**<https://hal.inria.fr/inria-00398182>**

Submitted on 24 Jun 2009

**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.

Fouille de données à l'aide de  
modèles stochastiques  
Segmentation temporo-spatiale des successions  
de cultures  
d'un territoire agricole à l'aide de HMM2

El Ghali Lazrak<sup>1</sup>    Marc Benoît<sup>1</sup>    Jean-François Mari<sup>1,2</sup>

<sup>1</sup> : INRA, UR 055, SAD ASTER  
domaine du Joly, F-88500 Mirecourt  
{lazrak, benoit}@mirecourt.inra.fr

<sup>2</sup> : UMR CNRS 7503 et INRIA-Grand Est, LORIA  
B.P. 239 F-54506 Vandœuvre-lès-Nancy  
jfmari@loria.fr

**Abstract**

We propose an original data mining method to segment agricultural landscapes based on a temporo spatial modelling of their crop successions. This method consists in (i) choosing the length of the crop succession, (ii) choosing the spatial resolution to sample the territory, (iii) specifying a conceptual model to represent the crop successions by means of second-order Hidden Markov Models, and (iv) finally segmenting the territory into homogeneous areas that will be further investigated.

Mots-clés : HMM2, segmentation temporo-spatiale, fouille de données  
keywords: HMM2, temporo-spatial segmentation, data mining

## 1 Introduction

Dans un contexte de développement durable, l'activité agricole se doit, entre autres exigences, de préserver le milieu naturel qu'elle utilise et transforme. Représenter la répartition de l'activité agricole dans l'espace à des échelles compatibles avec celles du déroulement d'enjeux environnementaux et/ou de processus écologiques va dans le sens des efforts visant à préserver et à rétablir l'équilibre dans les agro-écosystèmes [15].

L'étude des successions de couverts végétaux pour représenter l'activité agricole d'un agro-système a été initiée, depuis la fin du XIXe siècle,

dans le cadre de travaux sur les dynamiques de la végétation dans des écosystèmes naturels [8]. Les successions de cultures intègrent différentes logiques qui dictent le travail de l'agriculteur. La fouille des successions de cultures d'un territoire a pour objectif de mettre à jour ces logiques et de quantifier leurs importances sur l'évolution des paysages et l'impact des contraintes agricoles sur la biodiversité.

La méthode de fouille de données présentée dans cet article considère le territoire agricole comme une mosaïque de parcelles sur lesquelles se trouvent des couverts végétaux ou d'autres occupations telles que bâti, routes, ... La méthode consiste (i) à choisir la longueur de la succession temporelle d'occupations du sol (OCS), (ii) à choisir la résolution de l'échantillonnage spatial, (iii) à définir un modèle conceptuel pour la représentation des successions des OCS, et enfin, (iv) à créer une carte de paysages sous forme de zones homogènes vis-à-vis des successions d'occupation du sol et de croiser cette information avec la présence / absence des espèces animales considérées qui dégageront des zones qu'il faudra prospector plus profondément.

Après une introduction suivie de la présentation du matériel et des méthodes, la partie 3 décrit notre application dans laquelle nous recherchons le lien entre successions de cultures et présence d'une espèce animale protégée – le busard de Montagu – dans un territoire de  $350 \text{ Km}^2$  dans l'Ouest de la France. La conclusion esquisse la suite de ce travail préliminaire de fouille de données temporelles et spatiales.

## 2 Matériels et méthodes

### 2.1 Les occupations du sol

Le territoire agricole étudié –  $350 \text{ Km}^2$  dans la plaine céréalière de Niort – est enquêté depuis plus de 12 ans. La localisation et les occupations de ses parcelles sont relevées chaque année. Cette enquête est stockée dans un système d'information géographique (SIG) et est destinée à suivre les évolutions des occupations et des rotations de cultures notamment en ce qui concerne l'évolution des prairies : OCS essentielles pour la préservation de certaines espèces animales protégées. L'ensemble constitue un gros corpus d'informations temporelles et spatiales possédant un niveau de détails supérieur à ce qu'une analyse d'images satellites peut actuellement obtenir. Ce corpus est représenté sous la forme d'une matrice dans laquelle les colonnes représentent les OCS année par année et les lignes les différents sites enquêtés.

L'analyse des fréquences moyennes annuelles des OCS calculées sur les 12 années de la période d'étude fait ressortir 47 occupations du sol de la matrice de données. Les experts agronomes les regroupent ensuite en 10 catégories (*cf.* tableau 1) suivant une démarche tenant compte de la similitude des conduites culturales. Sont retenues les OCS : Blé (B), Tournesol (T), Colza (C), Urbain (U), Prairies et luzernes (P), Maïs (M), Forêts et friches (F), Orge d'hiver (O), raY-grass (Y), poiS (S) et Autres (A).

Catégorie d'occupation du sol	Occupation du sol	Fréquence cumulée
Blé (B)	blé, blé barbu, céréale	0.337
Tournesol (T)	tournesol, ray-grass suivi de tournesol	0.476
Colza (C)	Colza	0.600
Urbain (U)	bâti, péri-village, route	0.696
Prairies et Luzernes (P)	prairie permanente, prairie année 1, prairie temporaire (2-3 ans), prairie âge inconnu, luzerne 1 an, luzerne 2 ans, luzerne 3 ans, luzerne > 3 ans	0.6
Maïs (M)	maïs, ray-grass suivi de maïs	0.850
Forêts et friches (F)	forêt ou haie, friche	0.884
Orge d'hiver (O)	orge d'hiver	0.918
Ray-grass (Y)	ray-grass, ray-grass suivi de ray-grass	0.942
Pois (S)	Pois	0.964
Autres (A)	orge de printemps, vigne, jachère spontanée juin, moha, lin, avoine, trèfle, féverole, ray-grass suivi de labour, ray-grass suivi d'inconnu, jachère spontanée suivie de labour, mélange céréale légumineuse, culture printemps, moutarde, jardin / culture maraîchère, sorgho / millet, sorgho, millet, labour, tabac, autre culture	1.000

Table 1: Composition et fréquences moyennes des catégories d'occupation du sol adoptées

## 2.2 La mosaïque agricole

Le composant de base dans un territoire agricole est la parcelle : polygone de taille variable possédant un couvert – l'OCS – et éventuellement habitée par une espèce animale. Elle est délimitée géographiquement par une route, un chemin, ... ou artificiellement par l'agriculteur qui subdivise le territoire de son exploitation pour respecter un assolement : ensemble des surfaces allouées à chaque culture. Les frontières des parcelles changent chaque année en fonction des choix des agriculteurs (*cf.* Fig. 1). Pour tenir compte de ce changement, les enquêteurs définissent l'ensemble des micro-parcelles comme étant l'union de toutes les intersections de parcelles pendant la période d'étude. Il y a environ 20000 micro-parcelles dans le territoire étudié. Tous les points d'une micro-parcelle n'ont hébergé qu'une succession de cultures pendant la période d'étude.

L'analyse spatiale d'un territoire agricole vu comme une mosaïque de parcelles de taille quelconque présente des particularités dues aux caractéristiques des OCS. Dans l'analyse spatiale du territoire, la parcelle joue le rôle d'un pixel de taille variable possédant plusieurs modalités : les différentes OCS. L'OCS d'une parcelle à l'instant  $t$  s'inscrit dans une succession temporelle d'OCS qui intègre le savoir-faire de l'agriculteur qui met en valeur son territoire en fonction de différentes logiques. L'organisation dans le temps implique une organisation dans l'espace. Ainsi l'OCS d'une

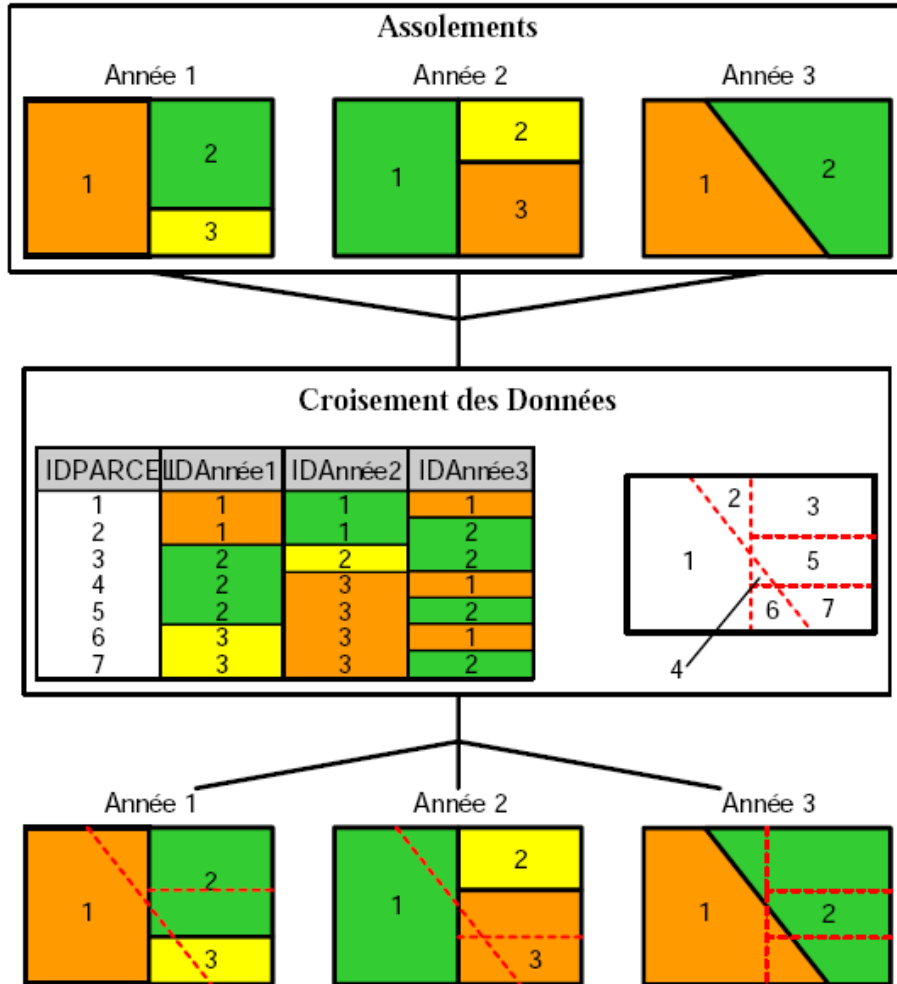


Figure 1: Exemple d'évolution des limites de parcelles pendant trois années successives. L'intersection des parcelles pendant cette période aboutit à la définition de sept micro-parcelles

parcelle une année donnée dépend de l'OCS les années précédentes sur cette parcelle ainsi que des OCS des parcelles voisines. Dans la mosaïque parcellaire, le système de voisinage est irrégulier. Une parcelle a un nombre quelconque de parcelles avec lesquelles elle partage une frontière commune. Nous modélisons la mosaïque parcellaire par un champ de Markov irrégulier dans lequel la parcelle s'inscrit dans une succession temporelle d'OCS. L'ensemble constitue une image de successions plutôt qu'une succession d'images et ne peut être étudié directement par les modèles numériques d'images appliqués au suivi de trajectoires d'objets [14]. Enfin, les territoires étudiés ne sont pas carrés et les parcelles ajoutées pour leur donner une forme carrée doivent être en nombre minimum afin de ne pas perturber la classification spatiale.

## 2.3 Modèles stochastiques pour représenter les successions de cultures de la mosaïque parcellaire

### 2.3.1 Modèles pour représenter l'évolution temporelle

Les successions de couverts végétaux se prêtent bien à la modélisation avec des modèles de Markov [11, 5]. Plus l'ordre du modèle de Markov est long, plus celui-ci peut modéliser des longues successions et plus il intéresse l'agronome. Les paramètres du modèle de Markov peuvent être estimés par expertise [5], comme ils peuvent être automatiquement calculés à partir d'un corpus d'observations en utilisant des algorithmes d'apprentissage comme celui de Baum-Welch [2] lorsqu'il s'agit de Modèles de Markov cachés (HMM). Dans notre cas, nous avons reproduit en adaptant la démarche de fouille effectuée sur des grands territoires en utilisant des HMM d'ordre 2 (HMM2) capables de modéliser [12, 4, 11] des rotations de cultures sur plus de deux ans, ce qui constitue un avantage sur les HMM d'ordre 1. Le HMM a l'avantage – comparé à la chaîne de Markov – de représenter l'état du processus par une distribution sur l'ensemble des observations plutôt que par une seule observation comme c'est le cas dans une chaîne de Markov. L'observation est habituellement une OCS ou une succession temporelle de 2 à 4 OCS. Ceci permet plus de souplesse. Enfin les HMM permettent une modélisation spatiale. Ce travail montre nos résultats en classification en utilisant des HMM2 aussi bien pour traiter les dimensions temporelle et spatiale des données ce qui donne plus de cohérence aux traitements.

### 2.3.2 Modèles pour représenter un champ spatial de Markov

Les champs de Markov permettent une classification non supervisée d'images numériques [10, 7, 1]. Pour contraindre la puissance de calcul nécessaire à leurs estimations, des méthodes approchées à l'aide de courbes fractales parcourant le plan et segmentées par un HMM ont été développées [3, 6]. Le modèle utilisé dans cet article pour segmenter l'espace utilise la dernière approche et s'appuie sur un HMM2 proposé [12, 11] dans le cadre de l'analyse des données spatio-temporelles *Ter-Uti* qui définissent un maillage spatial régulier de la France. Nous l'avons adapté pour traiter des pixels de taille différente - les micro-parcelles – dans un système de voisinage ir-

régulier. Nous utilisons une courbe fractale de Peano dont la “profondeur de fractalisation” s’ajuste à la taille de la micro-parcelle. Un processus récursif fusionne 4 points situés sur la courbe quand ils correspondent au motif générateur et qu’ils sont dans la même micro-parcelle. La figure 2 donne un exemple du parcours de la fractale dans une mosaïque parcelle échantillonnée par une image 16 x 16. Dans la suite des traitements spatiaux, le territoire sera représenté par la suite des points de la fractale.

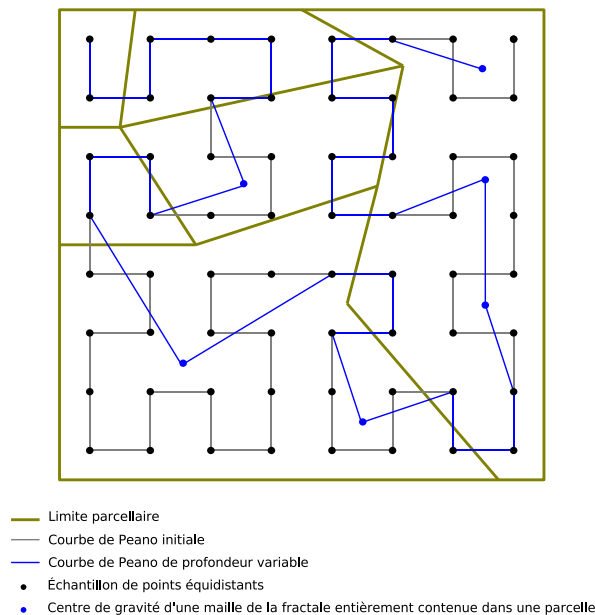


Figure 2: Prise en compte de la taille des parcelles dans le parcours du plan par la courbe fractale de Hilbert-Peano. L’exemple montre deux fusions successives dans la parcelle située en bas à gauche aboutissant à l’agglomération de 16 points en un seul

### 2.3.3 Modèles pour représenter un champ spatial de successions : les HMM hiérarchiques

Les HMM2 temporels et HMM2 spatiaux sont fusionnés dans un modèle de Markov caché hiérarchique (HHMM 2) [9, 13] et segmentent la représentation spatiale du territoire en zones homogènes, c’est à dire constituées de parcelles dont la succession d’OCS est modélisée par le même HMM2. La figure 3 donne un exemple d’HHMM2. L’ensemble des programmes permettant la spécification des HHMM2, leur apprentissage à l’aide d’un corpus de données, la segmentation en zones homogènes représentées par les états cachés de HMM2 ainsi que la création des shapefiles visualisant ces zones constituent la boîte à outils ARPENTAG (Analyse de Régularités dans

les Paysages : Environnement, Territoires et Agronomie), sur ensemble de CARROTAGE <sup>1</sup> utilisé pour l'analyse des successions d'OCS.

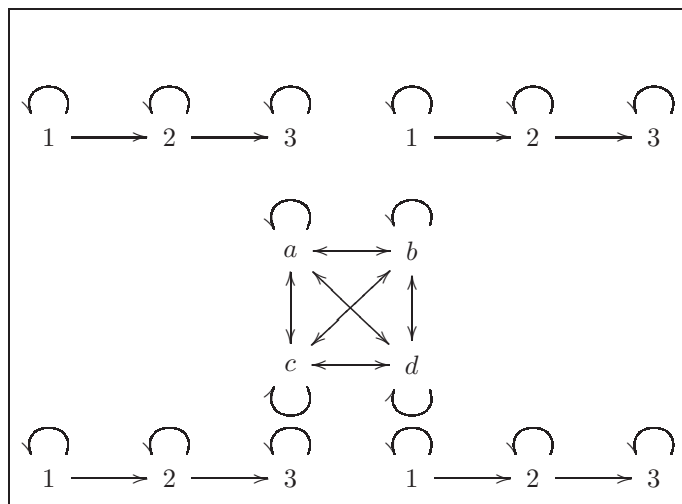


Figure 3: Chaque état a, b, c, d du HHMM2 est un HMM2 temporel dont les états sont 1, 2, 3.

### 3 Application

Notre but est d'identifier les régularités temporelles en termes de successions

d'OCS puis de les localiser dans le territoire afin de croiser les zones homogènes découvertes ("patches") avec la localisation des nids de busards de Montagu – espèce animale protégée – connue également des enquêteurs.

#### 3.1 Choix de la longueur des successions et de la résolution d'échantillonnage

La charge de calcul de l'estimation des paramètres des HMM peut être réduite en contrôlant deux facteurs: (i) la taille de la matrice des données à travers la résolution d'échantillonnage spatial, (ii) et la longueur de la succession d'occupation du sol.

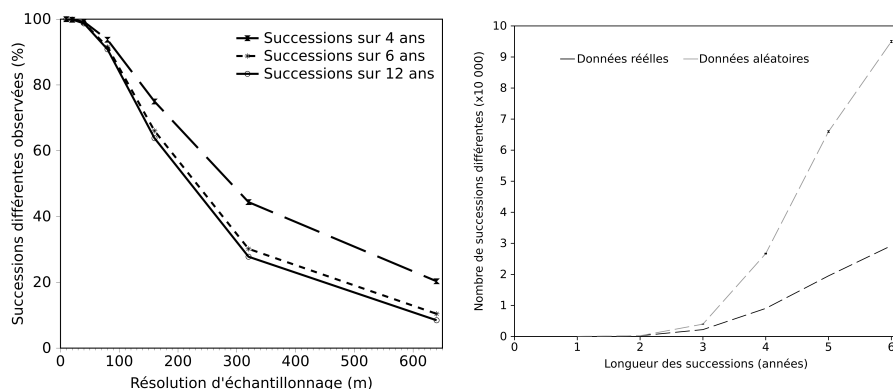
Afin de disposer d'un critère objectif pour le choix de la résolution spatiale, la perte d'information en terme de diversité des successions d'occupations du sol a été quantifiée pour des résolutions d'échantillonnage de moins en moins élevées (*cf.* Fig. 4(a)). On remarque qu'une résolution de 80 m x 80 m permet d'obtenir une matrice de données 64 fois plus réduite que la matrice de données initiale avec une perte de seulement 6% en termes de diversité d'information.

---

<sup>1</sup>licence GPL



Nous nous sommes intéressés ensuite à l'influence de la longueur des successions sur l'entropie du système. Nous avons comparé le nombre de successions différentes à celles produites aléatoirement dans le même territoire. Chaque OCS est tirée selon une loi uniforme sans tenir compte ni de sa localisation ni des OCS précédentes au même endroit. La figure 4(b) montre qu'à partir d'une longueur de succession de 4 années, la zone d'étude commence à se distinguer nettement du modèle où les successions sont choisies de manière aléatoire dans la zone d'étude.



(a) Nbr. de successions vs résolution spatiale (b) Nbr. de successions vs durée de la succession

Figure 4: (a) Évolution de la diversité spatiale en fonction de la résolution d'échantillonnage. (b) Comparaison de la diversité des successions d'occupations du sol entre les données réelles d'occupation du sol et des données générées aléatoirement pour différentes longueurs de successions

### 3.2 Influence de la résolution spatiale sur la longueur de la courbe fractale

Le SIG permet de définir une image raster en appliquant une grille de points régulièrement espacés. Le parcours de Hilbert-Peano sur une grille de points se détermine facilement quand la grille est un carré de côté  $2^n$ . Différentes tailles d'image – de 4096 à 256 correspondant à des résolutions spatiales de 10 à 160 mètres – ont été choisies pour contenir le territoire d'étude. Pour insérer le territoire d'étude dans une image carrée, nous l'entourons d'une "grande" micro-parcelle constituée de points "blancs" possédant l'occupation "Hors zone". La table 2 montre la réduction de la longueur de la fractale en fonction de la résolution spatiale. La colonne 4 donne le nombre de points de l'image situés sur le territoire.

On remarque qu'une résolution de 160 mètres ne procure aucun gain ; il n'y a pas eu d'agglomération de points à l'intérieur d'une même micro-parcelle sauf pour la micro-parcelle "blanche" constituée des points "Hors zone". La fractale passe par tous les points échantillonnés du territoire. Environ 95% des micro-parcelles ont moins 320 mètres de côté et donc

long. fractale	taille image	résolution (m)	nb. points échantillonnés
1160245	4096	10	4287149
481462	2048	20	1074298
182476	1024	40	269832
61285	512	80	68052
17830	256	160	17278

Table 2: Longueur de la fractale en fonction de la résolution spatiale adoptée pour représenter le territoire

échappent à tout regroupement de points. A partir d'une résolution de 40 mètres – 70% des micro-parcelles ont plus de 80 mètres de coté – l'intérêt du regroupement des points apparaît nettement. Un inconvénient de la fractalisation à profondeur variable est que le nombre de points ne rend plus compte facilement de la surface. Les distributions d'OCS dans une zone donnée ne sont plus des assolements.

### 3.3 Recherche des régularités temporelles

Nous représentons la diversité des successions d'occupations du sol en autant de classes qu'il y a de catégories d'occupation du sol. Les classes de successions sont notées  $S(X)$ ,  $X$  étant une catégorie d'OCS, et la classe  $S(X)$  se lit: classe des successions avec  $X$ . Le tableau 4 représente le modèle de recherche utilisé pour extraire les successions appartenant à une classe d'occupation donnée. Une classe d'occupation se présente donc sous forme d'une liste de quadruplets associés chacun à sa fréquence dans la matrice de données. Chaque quadruplet comporte, au moins dans un de ses 4 éléments constitutifs, une occurrence de la catégorie d'occupation du sol ayant donné son nom à la classe.

Dans le territoire d'étude, le Blé (B), le Tournesol (T) et le Colza (C) sont le plus souvent intégrés dans une même succession de 4 ans (p. ex. TBCB, TBTB, CBCB, TBBB, CBBB). Les fouilles de données avec une classe commune à ces trois cultures ont permis d'obtenir des résultats plus cohérents qu'avec les classes séparées. Cette classe est notée  $S(B,T,C)$  regroupant les quadruplets impliquant au moins l'une de ces trois cultures.

Année t	Année t+1	Année t+2	Année t+3
X	?	?	?
?	X	?	?
?	?	X	?
?	?	?	X

Table 4: Le motif de recherche pour l'extraction des quadruplets impliquant une catégorie  $X$  d'occupation du sol. ? représente une quelconque catégorie d'occupation du sol

### 3.4 Analyse temporelle

Suite aux résultats trouvés dans le paragraphe 3.1, nous avons utilisé une résolution spatiale de 80 mètres et des observations constituées de 4 OCS temporelles successives sur tous les points du territoire. Les régularités temporelles ont été recherchées par fouille de données à l'aide de HMM2 dont les états ont été initialisés à l'aide de distributions de quadruplets d'OCS correspondant aux motifs de recherche décrits dans le tableau 4. Le HMM2 possède 10 états : autant que de  $S(X)$ ,  $X$  étant une des 10 catégories d'OCS (cf. tableau 1). Tous les états sont inter connectés. Ce type de modèle correspond au modèle de "type 2" décrit dans [11]. Un

apprentissage de ce HMM2 par l'algorithme de Baum-Welch détermine les probabilités a posteriori de transitions entre états comme l'illustre la figure 5. Cette figure montre une monotonie paysagère en terme des successions d'occupations du sol. En effet, la largeur quasi constante des lignes horizontales et diagonales du diagramme indique que durant la période d'étude (1996-2007), aucun changement majeur n'a affecté la dynamique des successions d'occupation des sols dans la zone d'étude. Ceci nous a conduit, dans l'étape de spatialisation des régularités temporelles à considérer stationnaire le processus d'allocation des terres en terme des successions d'occupations du sol.

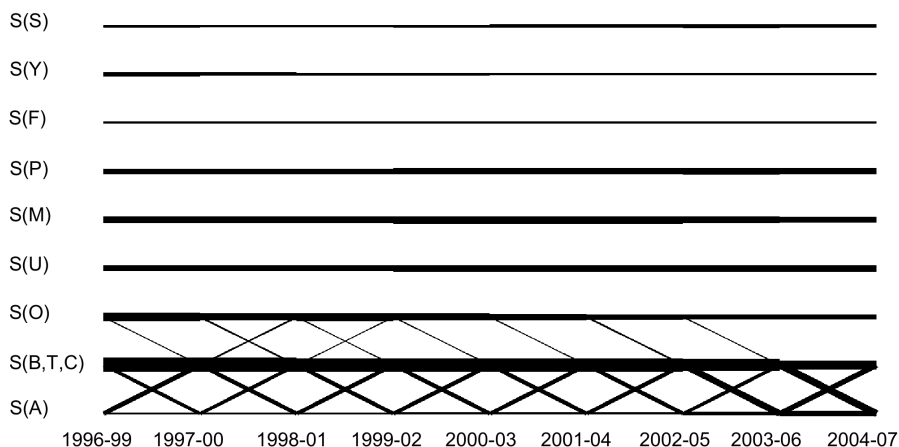


Figure 5: Probabilités a posteriori des transitions entre classes de successions d'occupation du sol dans un modèle de type 2. L'axe des abscisses représente la période d'étude répartie en sous périodes de 4 ans. L'axe des ordonnées représente les classes de successions d'occupations du sol impliquant: pois (S), raY-grass (Y), Forêts et friches (F), Prairies et luzernes (P), Maïs (M), Urbain (U), Orge d'hiver (O), Blé (B), Tournesol (T), Colza (C), Autres (A). L'épaisseur des lignes est proportionnelle à la fréquence des transitions. Les transitions horizontales indiquent une stabilité inter annuelle. Les transitions diagonales indiquent des changements inter annuels

### 3.5 Résultats de la classification temporo spatiale

Dans cette étape, nous localisons les régularités temporelles trouvées dans l'analyse temporelle. Comme le territoire ne présente aucune dynamique temporelle (*cf.* Fig. 5), chaque HMM2 temporel du HHMM2 ne contient qu'un seul état initialisé par les distributions trouvées dans l'analyse temporelle. La segmentation spatiale est faite par un HHMM2 ergodique (tous les états sont inter connectés comme le montre la Fig. 3). L'état possédant la plus forte probabilité a posteriori étiquette chaque point de la fractale. L'ensemble est rangé dans un shapefile qui représente la couche de classification telle qu'elle est visualisée dans la Fig. 6. Le croisement dans un

SIG de cette information avec la couche d'information relative à la localisation des nids de busards permet d'identifier des zones à prospecter plus profondément. Grâce à cette carte, l'enquêteur a une meilleure connaissance du type des exploitations agricoles rencontrées, ce qui va l'aider à mieux comprendre l'influence des mesures agro-environnementales (MAE) adoptées dans les exploitations qui visent en partie au maintien de ces espèces patrimoniales.

## 4 Conclusions

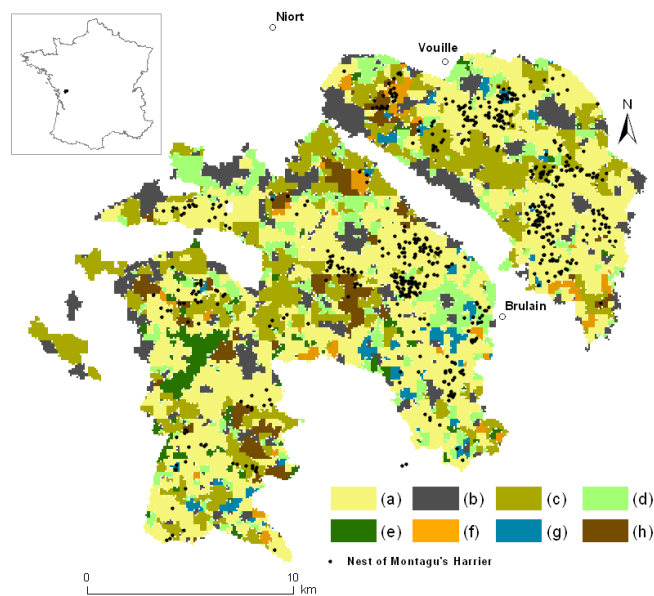
Cette méthode de fouille de données modélise la mosaïque parcellaire d'un territoire agricole à l'aide d'un champ de Markov caché temporo spatial irrégulier. Nous avons adopté un point de vue markovien pour représenter les évolutions temporelles et spatiales des occupations agricoles dans un paysage soumis à des enjeux environnementaux en utilisant des HMM2 capables de modéliser des dépendances sur une plus longue échelle temporelle et spatiale. En étudiant l'entropie du système d'allocation des terres, nous avons développé des méthodes pour fixer les paramètres temporels et spatiaux de ce champ. Le système de voisinage irrégulier de la mosaïque a été pris en compte par un parcours du plan par une courbe fractale dont la profondeur s'ajuste à la taille de la micro-parcelle. Cette information permet de ne représenter que par un point toutes les portions du territoire qui n'ont hébergé qu'une succession de cultures pendant la période d'étude. Cette modélisation a permis la localisation des régularités temporelles et spatiales dans les successions d'occupations du sol et de les croiser avec l'emplacement des nids de busards afin de mieux connaître les zones de prédilection de ces animaux. Nous envisageons maintenant de nous intéresser aux voisinages des parcelles afin d'identifier les quartiers de cultures plus aptes à accueillir le busard de Montagu et à fouiller leur évolution. Ce travail de fouille de données temporelles et spatiales répondant à des préoccupations écologiques et agronomiques est un travail long et multi disciplinaire qui fait collaborer des informaticiens qui adaptent et testent des modèles numériques de représentation de territoires, des agronomes qui cherchent à comprendre les logiques des agriculteurs à partir des résultats de la fouille et des écologues soucieux d'aménager l'habitat des espèces à protéger au sein de territoires agricoles soumis à des contraintes économiques.

## Remerciements

Nous remercions la région Lorraine, l'ANR BiodivAgrim, et l'API Ecoger pour leurs financements.

## References

- [1] S.A. Barker and P.J.W. Rayner. Unsupervised Image Segmentation Using Markov Random Field Models. *Pattern Recognition*, 33:587 – 602, 2000.



état	description des OCS
a	dominante de successions avec Tournesol, Blé et Colza (70%)
b	urbain et péri-urbain (60%)
c	dominante de successions avec Maïs (60%)
d	dominante de Prairies (50%)
e	dominante de Forêts (70%)
•	emplacement de nid de busard de Montagu

Figure 6: Carte des zones homogènes en termes des successions d'occupations du sol d'un territoire dans la Plaine de Niort durant la période 1996-2007. Les emplacements de nids de busards de Montagu sont représentés et déterminent des zones d'intérêt dans lesquelles les exploitations agricoles seront enquêtées pour évaluer l'influence des mesures agro-environnementales (MAE)

- [2] L. E. Baum. An Inequality and Associated Maximization Technique in Statistical Estimation for Probabilistic Functions of Markov Processes. *Inequalities*, 3:1 – 8, 1972.
- [3] B. Benmiloud and W. Pieczynski. Estimation des paramètres dans les chaînes de Markov cachés et segmentation d'images. *Traitement du signal*, 12(5):433 – 454, 95.
- [4] Marc Benoît, Florence Le Ber, Jean-Francois Mari, Catherine Mignolet, and Céline Schott. Carrotage, un logiciel pour la fouille de données agricoles. In *Colloque STIC et Environnement SE'2003, Rouen, France*. INSA Rouen, Jun 2003.
- [5] M.S. Castellazzi, G.A. Wood, P.J. Burgess, J. Morris, K.F. Conrad, and J.N. Perry. A systematic representation of crop rotations. *Agricultural Systems*, 97:26 – 33, 2008.
- [6] S. Derrode, C. Carincotte, and S. Bourennane. Unsupervised Image Segmentation Based on High-Order Markov Chains. In *IEEE Trans. ICASSP*, pages 769 – 772, 2004.
- [7] S. Geman and D. Geman. Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of Images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 6, 1984.
- [8] D.C. Glenn-Lewin and E. van der Maarel. *Plant Succession: Theory and Prediction*, chapter Patterns and processes of vegetation dynamics, pages 11 – 44. Chapman & Hall, London, 1992.
- [9] F.V. Jensen. *Bayesian Networks and Decision Graphs*. Springer, 2001.
- [10] Besag Julian. Statistical Analysis of Non-lattice Data. *The Statistician*, 24(3):179 – 195, 1975.
- [11] F. Le Ber, M. Benoit, C. Schott, J.-F. Mari, and C. Mignolet. Studying Crop Sequences With CarrotAge, a HMM-Based Data Mining Software. *Ecological Modelling*, 191(1):170 – 185, Jan 2006. <http://hal.archives-ouvertes.fr/hal-00017169/fr/>.
- [12] Jean-François Mari, Florence Le Ber, and Marc Benoît. Segmentation temporelle et spatiale de données agricoles. In *Actes des journées Cassini 2002, Crozon*, pages 251–272, septembre 2002.
- [13] Jean-françois Mari and Christine Largouet. *Raisonnements sur l'espace et le temps : des modèles aux applications*, chapter 9 et 12, pages 249 – 271 et 305 – 316. Lavoisier, <http://hal.archives-ouvertes.fr/hal-00126022/fr/> edition, mars 2007. ISBN : 2-7462-1455-5.
- [14] V. Noyon, C. Claramunt, and D. Devogele. A Relative Representation of Trajectories in Geographical Spaces. *Geoinformatica, Springer*, 14(11):479 – 496, 2007.
- [15] J.P. Deffontaines P. Thinon. Partage de l'espace rural pour la gestion des problèmes environnementaux et paysagers dans le vevin français. In *cah. Agr.*, number 8, pages 373 – 387, 1999.