

Extraction de connaissances agronomiques par fouille des voisinages entre occupations du sol

El Ghali Lazrak¹, Noémie Schaller², Jean-François Mari^{3,1}

¹INRA-SAD, F-88500 Mirecourt

²INRA/AgroParisTech, SAD-APT, F-78850 Thiverval-Grignon

³LORIA & INRIA-Grand Est F-54500 Vandœuvre-lès-Nancy

mardi 25 janvier, atelier Fouille de données complexes

La mosaïque parcellaire

données d'enquêtes du Centre d'études biologiques de Chizé (CEBC), Zone Atelier Plaine et Val de Sèvre

Definition (Parcelles,OCS)

- ▶ Le paysage agricole peut être perçu comme un assemblage de polygones de tailles différentes – les parcelles – où chaque parcelle porte une occupation du sol (OCS) ;
- ▶ Ne portent qu'un type de culture ;
- ▶ Re-découpées chaque année ;
- ▶ Organisées dans un système de voisinage irrégulier ;
- ▶ Les OCS sont choisies par les agriculteurs en fonction de différentes contraintes et opportunités ;

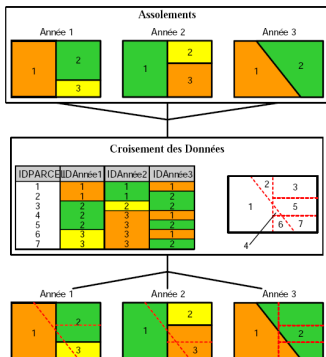
Corrélation spatiale et temporelle : on trouve fréquemment à proximité d'une parcelle l'OCS suivante que la parcelle portera.

La mosaïque parcellaire

données d'enquêtes du Centre d'études biologiques de Chizé (CEBC), Zone Atelier Plaine et Val de Sèvre

Definition (Micro parcelles)

- ▶ Les plus petits polygones (20 000) produits par la superposition des années.
- ▶ Leurs points ne portent qu'une succession de cultures (12 ans) ;



Les enquêtes sur le terrain (1992–2010)

Premier niveau d'échelle

- ▶ Effectuées chaque année ;
- ▶ Le territoire enquêté s'agrandit chaque année ;
- ▶ Les parcelles non enquêtées portent l'OCS **indéterminé**.

OCS	Fréq. cumulée
Blé (B)	0.337
Tournesol (T)	0.476
Colza (C)	0.600
Urbain (U)	0.696
Prairies et Luzernes(P)	0.774
Maïs (M)	0.850
Forêts et friches (F)	0.884
Orge d'hiver (O)	0.918
Ray-grass (R)	0.942
Pois (S)	0.964
Autres (A)	1.000

Table: Les modalités des OCS

Les enquêtes en exploitations agricoles (2006 – 2010)

deuxième niveau d'échelle

Les experts agronomes ont réalisé 67 enquêtes parmi les 185 exploitations ayant toute leur surface dans la zone d'étude pour **comprendre le fonctionnement** global de l'exploitation agricole.

- ▶ 22 enquêtes en 2006 et 19 enquêtes en 2007 ont porté respectivement sur les stratégies des éleveurs et des agriculteurs pour faire face aux **sécheresses estivales** et aux interdictions d'**irrigation** ;
- ▶ 12 enquêtes en 2009 ont porté sur les décisions des agriculteurs relatives à l'**allocation des cultures** dans les parcelles et au **découpage des parcelles** ;
- ▶ 14 enquêtes en 2010 ont porté sur l'évolution dans le temps des **assolements** annuels choisis par l'exploitant ;
- ▶ Les enquêtes étaient semi-structurées pour encourager l'agriculteur à expliciter les raisons de ses choix et leur évolution au fil du temps, notamment la façon d'allouer les successions de cultures dans les parcelles.

Les enquêtes en exploitations agricoles (2006 – 2010)

deuxième niveau d'échelle

Les experts agronomes ont réalisé 67 enquêtes parmi les 185 exploitations ayant toute leur surface dans la zone d'étude pour **comprendre le fonctionnement** global de l'exploitation agricole.

- ▶ 22 enquêtes en 2006 et 19 enquêtes en 2007 ont porté respectivement sur les stratégies des éleveurs et des agriculteurs pour faire face aux **sécheresses estivales** et aux interdictions d'**irrigation** ;
- ▶ 12 enquêtes en 2009 ont porté sur les décisions des agriculteurs relatives à l'**allocation des cultures** dans les parcelles et au **découpage des parcelles** ;
- ▶ 14 enquêtes en 2010 ont porté sur l'évolution dans le temps des **assolements** annuels choisis par l'exploitant ;
- ▶ Les enquêtes étaient semi-structurées pour encourager l'agriculteur à expliciter les raisons de ses choix et leur évolution au fil du temps, notamment la façon d'allouer les successions de cultures dans les parcelles.

Les enquêtes en exploitations agricoles (2006 – 2010)

deuxième niveau d'échelle

Les experts agronomes ont réalisé 67 enquêtes parmi les 185 exploitations ayant toute leur surface dans la zone d'étude pour **comprendre le fonctionnement** global de l'exploitation agricole.

- ▶ 22 enquêtes en 2006 et 19 enquêtes en 2007 ont porté respectivement sur les stratégies des éleveurs et des agriculteurs pour faire face aux **sécheresses estivales** et aux interdictions d'**irrigation** ;
- ▶ 12 enquêtes en 2009 ont porté sur les décisions des agriculteurs relatives à l'**allocation des cultures** dans les parcelles et au **découpage des parcelles** ;
- ▶ 14 enquêtes en 2010 ont porté sur l'évolution dans le temps des **assolements** annuels choisis par l'exploitant ;
- ▶ Les enquêtes étaient semi-structurées pour encourager l'agriculteur à expliciter les raisons de ses choix et leur évolution au fil du temps, notamment la façon d'allouer les successions de cultures dans les parcelles.

Les enquêtes en exploitations agricoles (2006 – 2010)

deuxième niveau d'échelle

Les experts agronomes ont réalisé 67 enquêtes parmi les 185 exploitations ayant toute leur surface dans la zone d'étude pour [comprendre le fonctionnement](#) global de l'exploitation agricole.

- ▶ 22 enquêtes en 2006 et 19 enquêtes en 2007 ont porté respectivement sur les stratégies des éleveurs et des agriculteurs pour faire face aux **sécheresses estivales** et aux interdictions d'**irrigation** ;
- ▶ 12 enquêtes en 2009 ont porté sur les décisions des agriculteurs relatives à l'**allocation des cultures** dans les parcelles et au **découpage des parcelles** ;
- ▶ 14 enquêtes en 2010 ont porté sur l'évolution dans le temps des **assolements** annuels choisis par l'exploitant ;
- ▶ Les enquêtes étaient semi-structurées pour encourager l'agriculteur à expliciter les raisons de ses choix et leur évolution au fil du temps, notamment la façon d'allouer les successions de cultures dans les parcelles.

Modéliser les décisions des agriculteurs grâce à des enquêtes

Deux modèles agronomiques pour construire des enquêtes

Le fonctionnement de l'exploitation agricole est évalué par des enquêtes faites dans les exploitations agricole et construites à l'aide de deux modèles :

- ▶ modèle pour l'action [Sebillotte and Soler(1990)];
- ▶ modèle d'utilisation des ressources dans l'exploitation [Aubry et al.(1998)Aubry, Biarnes, Maxime, and Papy]

L'expert agronome construit un cadre conceptuel générique pour modéliser les décisions des agriculteurs à travers

- ▶ les variables de décision ;
- ▶ les déterminants ;
- ▶ les règles de décisions.

Modéliser les décisions des agriculteurs grâce à des enquêtes

Les variables de décision donnent une réponse à la question : "En quoi consiste la décision ?"

zone cultivable de la culture = l'ensemble des parcelles adaptées à cette culture,

taille de la sole = surface totale d'une culture une année donnée sur l'exploitation,

délai de retour = temps minimum à attendre avant de replanter la même culture sur la même parcelle ;

les couples de cultures précédent / suivant acceptables
= succession de culture

Les déterminants influencent les variables de décision.

Les règles de décision (si .. alors .. sinon ..) définies et suivies par un agriculteur, en fonction des déterminants, pour faire son choix et donner une valeur à chacune des variables de décision.

Modéliser les décisions des agriculteurs grâce à des enquêtes

Les variables de décision donnent une réponse à la question : "En quoi consiste la décision ?"

Les déterminants influencent les variables de décision.

- ▶ quantitatifs,
- ▶ qualitatifs,
- ▶ internes (par exemple les ressources de l'exploitation agricole)
- ▶ externes à l'exploitation agricole (par exemple les conditions du marché, le climat, ...).

Les règles de décision (si .. alors .. sinon ..) définies et suivies par un agriculteur, en fonction des déterminants, pour faire son choix et donner une valeur à chacune des variables de décision.

Modéliser les décisions des agriculteurs grâce à des enquêtes

Les variables de décision donnent une réponse à la question : "En quoi consiste la décision ?"

Les déterminants influencent les variables de décision.

Les règles de décision (si .. alors .. sinon ..) définies et suivies par un agriculteur, en fonction des déterminants, pour faire son choix et donner une valeur à chacune des variables de décision.

Les objectifs de la fouille

- ▶ Comprendre l'évolution de la mosaïque parcellaire dans le temps ;
- ▶ Evaluer l'impact des décisions des agriculteurs dans le paysage ;
- ▶ Trouver les modifications qui s'expliquent au niveau des exploitations agricoles (= dégâts collatéraux ? “on ne l'a pas décidé, mais ça s'est produit”) ;
- ▶ Utiliser deux approches d'extraction de connaissances :
 - approche descendante (déductive)** à base d'enquêtes ;
 - approche ascendante (inductive)** qui recherche des régularités stochastiques dans les données pour les expliquer.
- ▶ Analyser les dépendances temporelles et spatiales entre OCS à partir de 2 hypothèses :
 - hypothèse de champ de Markov** : l'OCS d'une parcelle dépend de l'OCS des parcelles voisines ;
 - hypothèse de chaîne de Markov** : l'OCS d'une parcelle une année donnée dépend des OCS trouvées sur cette parcelle les années précédentes.

Les objectifs de la fouille

- ▶ Comprendre l'évolution de la mosaïque parcellaire dans le temps ;
- ▶ Evaluer l'impact des décisions des agriculteurs dans le paysage ;
- ▶ Trouver les modifications qui s'expliquent au niveau des exploitations agricoles (= dégâts collatéraux ? “on ne l'a pas décidé, mais ça s'est produit”) ;
- ▶ Utiliser deux approches d'extraction de connaissances :
 - approche descendante (déductive)** à base d'enquêtes ;
 - approche ascendante (inductive)** qui recherche des régularités stochastiques dans les données pour les expliquer.
- ▶ Analyser les dépendances temporelles et spatiales entre OCS à partir de 2 hypothèses :
 - hypothèse de champ de Markov** : l'OCS d'une parcelle dépend de l'OCS des parcelles voisines ;
 - hypothèse de chaîne de Markov** : l'OCS d'une parcelle une année donnée dépend des OCS trouvées sur cette parcelle les années précédentes.

Les objectifs de la fouille

- ▶ Comprendre l'évolution de la mosaïque parcellaire dans le temps ;
- ▶ Evaluer l'impact des décisions des agriculteurs dans le paysage ;
- ▶ Trouver les modifications qui s'expliquent au niveau des exploitations agricoles (= dégâts collatéraux ? “on ne l'a pas décidé, mais ça s'est produit”) ;
- ▶ Utiliser deux approches d'extraction de connaissances :
 - approche descendante (déductive)** à base d'enquêtes ;
 - approche ascendante (inductive)** qui recherche des régularités stochastiques dans les données pour les expliquer.
- ▶ Analyser les dépendances temporelles et spatiales entre OCS à partir de 2 hypothèses :
 - hypothèse de champ de Markov** : l'OCS d'une parcelle dépend de l'OCS des parcelles voisines ;
 - hypothèse de chaîne de Markov** : l'OCS d'une parcelle une année donnée dépend des OCS trouvées sur cette parcelle les années précédentes.

Difficultés liées aux données issues de la mosaïque

On s'intéresse aux 4 voisins d'ordre 1 d'un site S au temps t :

- ▶ Soient No_t, So_t, Es_t, We_t , les 4 occupations à différentes distances au Nord, Sud, Est et Ouest de S_t .
- ▶ On étudie la distribution du quintuplet $S_t, No_t, So_t, Es_t, We_t$, $t = 1, 12$.
- ▶ Le voisinage est irrégulier et variable dans le temps, on applique une grille d'échantillonnage spatial constante dans le temps
- ▶ Le territoire d'étude s'est étendu au fil du temps, certains points ont été étiquetés *indéterminé* avant de l'être par une OCS.
- ▶ Corpus de grande dimension ;

Difficultés liées aux données issues de la mosaïque

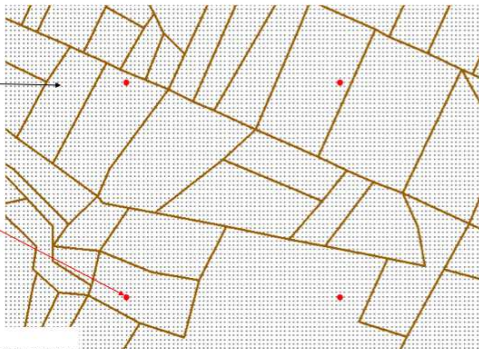
On s'intéresse aux 4 voisins d'ordre 1 d'un site S au temps t :

- ▶ Soient No_t, So_t, Es_t, We_t , les 4 occupations à différentes distances au Nord, Sud, Est et Ouest de S_t .
- ▶ On étudie la distribution du quintuplet $S_t, No_t, So_t, Es_t, We_t$, $t = 1, 12$.
- ▶ Le voisinage est irrégulier et variable dans le temps, on applique une grille d'échantillonnage spatiale constante dans le temps

Avec une résolution d'échantillonnage fine, la grande taille du corpus alourdi la charge de calcul.

Avec une résolution d'échantillonnage grossière, les plus petites parcelles sont omises.

 Limite parcellaire



Difficultés liées aux données issues de la mosaïque

On s'intéresse aux 4 voisins d'ordre 1 d'un site S au temps t :

- ▶ Soient No_t, So_t, Es_t, We_t , les 4 occupations à différentes distances au Nord, Sud, Est et Ouest de S_t .
 - ▶ On étudie la distribution du quintuplet $S_t, No_t, So_t, Es_t, We_t$, $t = 1, 12$.
 - ▶ Le voisinage est irrégulier et variable dans le temps, on applique une grille d'échantillonnage spatial constante dans le temps
 - ▶ Le territoire d'étude s'est étendu au fil du temps, certains points ont été étiquetés *indéterminé* avant de l'être par une OCS.
- ▶ Corpus de grande dimension ;

Difficultés liées aux données issues de la mosaïque

On s'intéresse aux 4 voisins d'ordre 1 d'un site S au temps t :

- ▶ Soient No_t, So_t, Es_t, We_t , les 4 occupations à différentes distances au Nord, Sud, Est et Ouest de S_t .
- ▶ On étudie la distribution du quintuplet $S_t, No_t, So_t, Es_t, We_t$, $t = 1, 12$.
- ▶ Le voisinage est irrégulier et variable dans le temps, on applique une grille d'échantillonnage spatial constante dans le temps
- ▶ Le territoire d'étude s'est étendu au fil du temps, certains points ont été étiquetés *indéterminé* avant de l'être par une OCS.
- ▶ Corpus de grande dimension ;

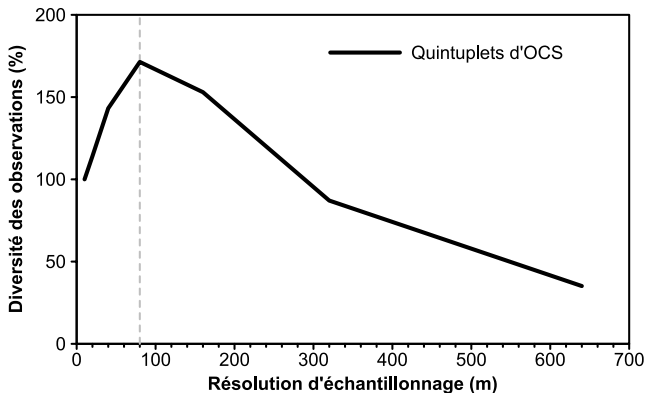


Figure: Nombre de quintuplés (la configuration d'un site augmenté de ses 4 voisins) suivant la résolution spatiale. Le nombre de quintuplés à 10 m est considéré comme référence (100%). La courbe montre les mêmes propriétés qu'en analyse de textures d'images numériques

Nettoyage des données

Segmentation temporelle à l'aide de HMM2

but :

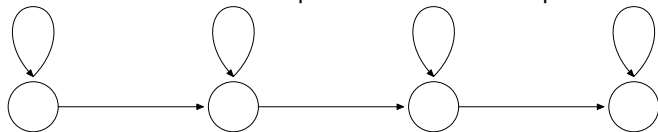
- ▶ éliminer le biais introduit par l'OCS indéterminé ;
- ▶ construction d'un *HMM2* pour isoler les sous séquences *indéterminé* ;
- ▶ apprentissage sur le corpus de quintuplets ;

Nettoyage des données

Segmentation temporelle à l'aide de HMM2

but :

- ▶ éliminer le biais introduit par l'OCS indéterminé ;
- ▶ construction d'un *HMM2* pour isoler les sous séquences *indéterminé* ;



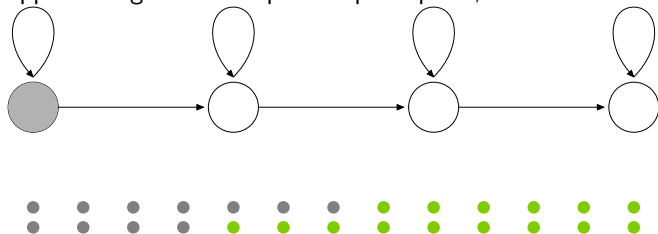
- ▶ apprentissage sur le corpus de quintuplets ;

Nettoyage des données

Segmentation temporelle à l'aide de HMM2

but :

- ▶ éliminer le biais introduit par l'OCS indéterminé ;
- ▶ construction d'un *HMM2* pour isoler les sous séquences *indéterminé* ;
- ▶ apprentissage sur le corpus de quintuplets ;

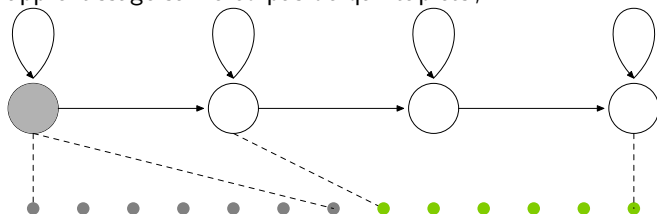


Nettoyage des données

Segmentation temporelle à l'aide de HMM2

but :

- ▶ éliminer le biais introduit par l'OCS indéterminé ;
- ▶ construction d'un *HMM2* pour isoler les sous séquences *indéterminé* ;
- ▶ apprentissage sur le corpus de quintuplets ;



Le *HMM2* aligne les observations d'une façon élastique ;

Segmenter les 12 années en périodes stationnaires

but : Isoler deux périodes : début / fin encadrant la sécheresse de 2006.

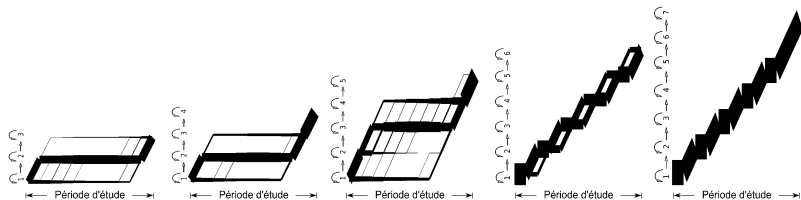


Figure: L'épaisseur des traits est proportionnelle à la probabilité *a posteriori* des états du *HMM2*. La segmentation recherchée comporte au moins 2 périodes de plusieurs années ne se recouvrant pas et ne contenant pas l'OCS *indéterminé*. Dans un *HMM2* de 6 états, les états 2 et 5 identifient deux périodes disjointes

Réduire la dimension des données

Contenu d'un état du HMM2 : extrait des voisinages du blé

0	...	les quintuplets "plein champs"
9	0.0069	blé, blé, blé, tourn., blé
10	0.0059	blé, blé, blé, blé,tourn.
11	0.0053	blé, blé, tourn., blé, blé
13	0.0050	blé, tourn., blé, blé, blé
17	0.0038	blé, blé, tourn., tourn., blé
18	0.0037	blé, tourn., blé, blé, tourn.
27	0.0028	blé, blé, tourn., blé, tourn.
32	0.0027	blé, tourn., blé, tourn., blé
...
999

Réduire la dimension des données

Contenu d'un état du HMM2 : extrait des voisinages du blé

0
9	0.0069	blé, blé, blé, tourn., blé
10	0.0059	blé, blé, blé, blé, tourn.
11	0.0053	blé, blé, tourn., blé, blé
13	0.0050	blé, tourn., blé, blé, blé
17	0.0038	blé, blé, tourn., tourn., blé
18	0.0037	blé, tourn., blé, blé, tourn.
27	0.0028	blé, blé, tourn., blé, tourn.
32	0.0027	blé, tourn., blé, tourn., blé
...
999

Le tournesol apparaît simultanément dans toutes les directions. Isotropie du territoire ??

Réduire la dimension des données

Isotropie du territoire

- ▶ $P(V, S)$ représente la probabilité d'avoir le site actuel occupé par l'OCS y ($S = y$) et d'avoir l'OCS x sur le site voisin ($V = x$)
- ▶ On montre que $P(S, V)$ est la même quelle que soit la direction de voisinages – No, So, Es, We –
- ▶ La distance entre deux distributions est calculée à l'aide de la divergence :

$$\text{div}(f, g) = \frac{1}{2} \sum_x (f(x) - g(x)) \log \frac{f(x)}{g(x)} \quad (1)$$

quand f et g sont deux distributions discrètes sur le même espace décrit par x .

- ▶ On utilisera des cliques (2 points voisins) plutôt que des quintuplets.

Réduire la dimension des données

Isotropie du territoire

- ▶ $P(V, S)$ représente la probabilité d'avoir le site actuel occupé par l'OCS y ($S = y$) et d'avoir l'OCS x sur le site voisin ($V = x$)
- ▶ On montre que $P(S, V)$ est la même quelle que soit la direction de voisinages – No, So, Es, We –
- ▶ La distance entre deux distributions est calculée à l'aide de la divergence :

$$\text{div}(f, g) = \frac{1}{2} \sum_x (f(x) - g(x)) \log \frac{f(x)}{g(x)} \quad (1)$$

quand f et g sont deux distributions discrètes sur le même espace décrit par x .

- ▶ On utilisera des cliques (2 points voisins) plutôt que des quintuplets.

Réduire la dimension des données

Isotropie du territoire

- ▶ $P(V, S)$ représente la probabilité d'avoir le site actuel occupé par l'OCS y ($S = y$) et d'avoir l'OCS x sur le site voisin ($V = x$)
- ▶ On montre que $P(S, V)$ est la même quelle que soit la direction de voisinages – No, So, Es, We –
- ▶ La distance entre deux distributions est calculée à l'aide de la divergence :

$$\text{div}(f, g) = \frac{1}{2} \sum_x (f(x) - g(x)) \log \frac{f(x)}{g(x)} \quad (1)$$

quand f et g sont deux distributions discrètes sur le même espace décrit par x .

- ▶ On utilisera des cliques (2 points voisins) plutôt que des quintuplets.

Réduire la dimension des données

Isotropie du territoire

- ▶ $P(V, S)$ représente la probabilité d'avoir le site actuel occupé par l'OCS y ($S = y$) et d'avoir l'OCS x sur le site voisin ($V = x$)
- ▶ On montre que $P(S, V)$ est la même quelle que soit la direction de voisinages – No, So, Es, We –
- ▶ La distance entre deux distributions est calculée à l'aide de la divergence :

$$\text{div}(f, g) = \frac{1}{2} \sum_x (f(x) - g(x)) \log \frac{f(x)}{g(x)} \quad (1)$$

quand f et g sont deux distributions discrètes sur le même espace décrit par x .

- ▶ On utilisera des **cliques** (2 points voisins) plutôt que des quintuplets.

mesure de co-localisation

$$PMI(x, y) = \log\left(\frac{P(V = x/S = y)}{P(V = x)}\right) = \log\left(\frac{P(V = x, S = y)}{P(V = x) \times P(S = y)}\right)$$

- ▶ $P(V/S)$ n'est pas une bonne mesure pour évaluer la co-localisation de deux DCS car elle dépend des probabilités marginales des DCS.
- ▶ Nous utilisons l'information mutuelle spécifique (PMI comme Pointwise Mutual Information) qui représente l'information apportée par la connaissance d'une variable sur l'autre.
- ▶ Une valeur positive signifie que le couple d'DCS (x, y) est co-localisé : les DCS x et y s'attirent.
- ▶ Une valeur négative signifie que les DCS se repoussent,
- ▶ l'agronome peut dans ce cas expliquer ou rechercher par enquêtes auprès des agriculteurs la (les) raison(s) de placer ou d'éviter de mettre ces DCS côte à côte.

mesure de co-localisation

$$PMI(x, y) = \log\left(\frac{P(V = x/S = y)}{P(V = x)}\right) = \log\left(\frac{P(V = x, S = y)}{P(V = x) \times P(S = y)}\right)$$

- ▶ $P(V/S)$ n'est pas une bonne mesure pour évaluer la co-localisation de deux DCS car elle dépend des probabilités marginales des DCS.
- ▶ Nous utilisons l'information mutuelle spécifique (PMI comme Pointwise Mutual Information) qui représente l'information apportée par la connaissance d'une variable sur l'autre.
- ▶ Une valeur positive signifie que le couple d'DCS (x, y) est co-localisé : les DCS x et y s'attirent.
- ▶ Une valeur négative signifie que les DCS se repoussent,
- ▶ l'agronome peut dans ce cas expliquer ou rechercher par enquêtes auprès des agriculteurs la (les) raison(s) de placer ou d'éviter de mettre ces DCS côte à côte.

mesure de co-localisation

$$PMI(x, y) = \log\left(\frac{P(V = x/S = y)}{P(V = x)}\right) = \log\left(\frac{P(V = x, S = y)}{P(V = x) \times P(S = y)}\right)$$

- ▶ $P(V/S)$ n'est pas une bonne mesure pour évaluer la co-localisation de deux OCS car elle dépend des probabilités marginales des OCS.
- ▶ Nous utilisons l'information mutuelle spécifique (PMI comme Pointwise Mutual Information) qui représente l'information apportée par la connaissance d'une variable sur l'autre.
- ▶ Une valeur positive signifie que le couple d'OCS (x, y) est co-localisé : les OCS x et y s'attirent.
- ▶ Une valeur négative signifie que les OCS se repoussent,
- ▶ l'agronome peut dans ce cas expliquer ou rechercher par enquêtes auprès des agriculteurs la (les) raison(s) de placer ou d'éviter de mettre ces OCS côte à côte.

mesure de co-localisation

$$PMI(x, y) = \log\left(\frac{P(V = x/S = y)}{P(V = x)}\right) = \log\left(\frac{P(V = x, S = y)}{P(V = x) \times P(S = y)}\right)$$

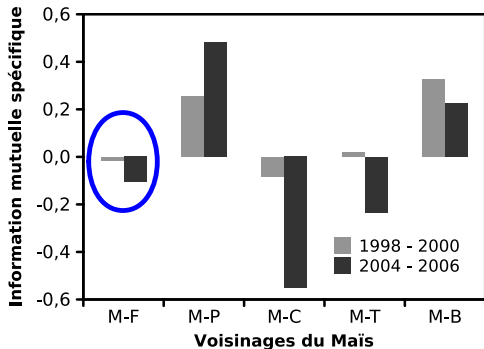
- ▶ $P(V/S)$ n'est pas une bonne mesure pour évaluer la co-localisation de deux OCS car elle dépend des probabilités marginales des OCS.
- ▶ Nous utilisons l'information mutuelle spécifique (PMI comme Pointwise Mutual Information) qui représente l'information apportée par la connaissance d'une variable sur l'autre.
- ▶ Une valeur positive signifie que le couple d'OCS (x, y) est co-localisé : les OCS x et y s'attirent.
- ▶ Une valeur négative signifie que les OCS se repoussent,
- ▶ l'agronome peut dans ce cas expliquer ou rechercher par enquêtes auprès des agriculteurs la (les) raison(s) de placer ou d'éviter de mettre ces OCS côte à côte.

mesure de co-localisation

$$PMI(x, y) = \log\left(\frac{P(V = x/S = y)}{P(V = x)}\right) = \log\left(\frac{P(V = x, S = y)}{P(V = x) \times P(S = y)}\right)$$

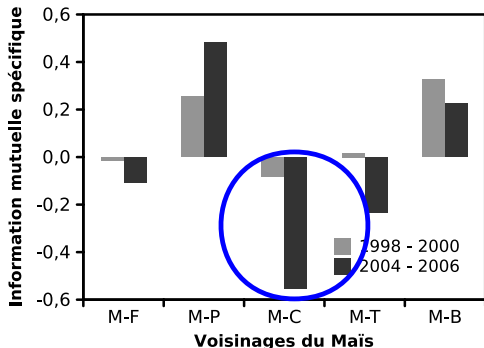
- ▶ $P(V/S)$ n'est pas une bonne mesure pour évaluer la co-localisation de deux OCS car elle dépend des probabilités marginales des OCS.
- ▶ Nous utilisons l'information mutuelle spécifique (PMI comme Pointwise Mutual Information) qui représente l'information apportée par la connaissance d'une variable sur l'autre.
- ▶ Une valeur positive signifie que le couple d'OCS (x, y) est co-localisé : les OCS x et y s'attirent.
- ▶ Une valeur négative signifie que les OCS se repoussent,
- ▶ l'agronome peut dans ce cas expliquer ou rechercher par enquêtes auprès des agriculteurs la (les) raison(s) de placer ou d'éviter de mettre ces OCS côte à côte.

Impacts des modifications d'assolement sur le paysage



Suite à la diminution des surfaces allouées au maïs, celui ci s'éloigne d'abord des forêts qui abritent les ravageurs ;

Impacts des modifications d'assolement sur le paysage

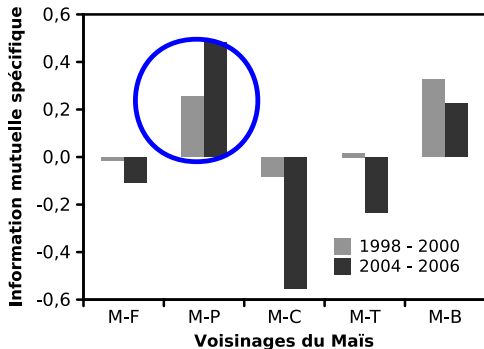


Suite à la sécheresse, le maïs disparaît du voisinage des terres "à colza" (C) = terres de plaine, séchantes, et peu adaptées à la culture de maïs sans irrigation non restreinte ;

Explication de régularités stochastiques

Une nouvelle conduite du maïs

- ▶ Rien n'indiquait dans les enquêtes que les maïs et les prairies devaient devenir voisins;

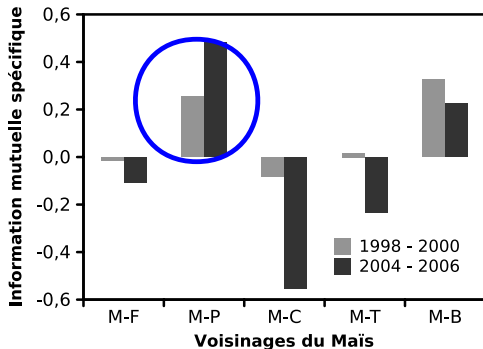


- ▶ les agriculteurs réduisent les surfaces en maïs en raison des risques de sécheresse estivale et le concentrent dans les terrains les plus humides = terres "à prairies".
- ▶ Les éleveurs étendent même la surface des prairies pour sécuriser la production de fourrages dans le cas où la production de maïs serait insuffisante, d'où la co-localisation maïs - prairies.

Explication de régularités stochastiques

Une nouvelle conduite du maïs

- ▶ Rien n'indiquait dans les enquêtes que les maïs et les prairies devaient devenir voisins;

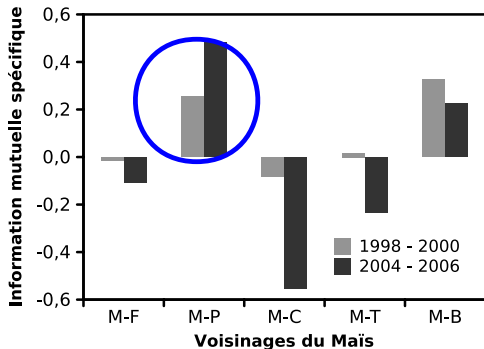


- ▶ les agriculteurs réduisent les surfaces en maïs en raison des risques de sécheresse estivale et le concentrent dans les terrains les plus humides = terres "à prairies".
- ▶ Les éleveurs étendent même la surface des prairies pour sécuriser la production de fourrages dans le cas où la production de maïs serait insuffisante, d'où la co-localisation maïs - prairies.

Explication de régularités stochastiques

Une nouvelle conduite du maïs

- ▶ Rien n'indiquait dans les enquêtes que les maïs et les prairies devaient devenir voisins ;



- ▶ les agriculteurs réduisent les surfaces en maïs en raison des risques de sécheresse estivale et le concentrent dans les terrains les plus humides = terres "à prairies".
- ▶ Les éleveurs étendent même la surface des prairies pour sécuriser la production de fourrages dans le cas où la production de maïs serait insuffisante, d'où la co-localisation maïs - prairies.

Conclusions

- ▶ Les décisions des agriculteurs, bien que prises à un niveau individuel et à une autre échelle, peuvent être génériques et entraîner des régularités dans le paysage ;
- ▶ Inversement, des régularités dans le paysage révèlent des processus de conduite d'exploitation plus complexes.
- ▶ L'extraction de connaissances par enquêtes en exploitations agricoles et par fouille de régularités spatiales au niveau de l'organisation des parcelles permet une analyse plus précise des résultats ;
- ▶ Intérêt en simulation de paysages à partir de scénarios pour tester des nouvelles conduites d'exploitation.

Conclusions

- ▶ Les décisions des agriculteurs, bien que prises à un niveau individuel et à une autre échelle, peuvent être génériques et entraîner des régularités dans le paysage ;
- ▶ Inversement, des régularités dans le paysage révèlent des processus de conduite d'exploitation plus complexes.
- ▶ L'extraction de connaissances par enquêtes en exploitations agricoles et par fouille de régularités spatiales au niveau de l'organisation des parcelles permet une analyse plus précise des résultats ;
- ▶ Intérêt en simulation de paysages à partir de scénarios pour tester des nouvelles conduites d'exploitation.

Conclusions

- ▶ Les décisions des agriculteurs, bien que prises à un niveau individuel et à une autre échelle, peuvent être génériques et entraîner des régularités dans le paysage ;
- ▶ Inversement, des régularités dans le paysage révèlent des processus de conduite d'exploitation plus complexes.
- ▶ L'extraction de connaissances par enquêtes en exploitations agricoles et par fouille de régularités spatiales au niveau de l'organisation des parcelles permet une analyse plus précise des résultats ;
- ▶ Intérêt en simulation de paysages à partir de scénarios pour tester des nouvelles conduites d'exploitation.

Conclusions

- ▶ Les décisions des agriculteurs, bien que prises à un niveau individuel et à une autre échelle, peuvent être génériques et entraîner des régularités dans le paysage ;
- ▶ Inversement, des régularités dans le paysage révèlent des processus de conduite d'exploitation plus complexes.
- ▶ L'extraction de connaissances par enquêtes en exploitations agricoles et par fouille de régularités spatiales au niveau de l'organisation des parcelles permet une analyse plus précise des résultats ;
- ▶ Intérêt en simulation de paysages à partir de scénarios pour tester des nouvelles conduites d'exploitation.



C. Aubry, A. Biarnes, F. Maxime, and F. Papy.

Modélisation de l'organisation technique de la production dans l'exploitation agricole : la constitution de système de culture.

Etud Rech Syst Agraires Dév, (31) :25–43, 1998.



N. Schaller, C. Aubry, A. Havet, and P. Martin.

Diversity of farmers' adaptation strategies in a context of changes and consequences on land-use dynamics : a methodological approach.

In Proceedings of the 1st Latin American and European congress on co-innovation of sustainable rural livelihood systems (Eulacias project), pages 189–192, Uruguay, 2010.



M. Sebillotte and Soler.

Modélisation systémique et systèmes agraires, chapter Les processus de décision des agriculteurs : acquis et questions vives, pages 93–102.

INRA Paris, 1990.