



Soutenance de thèse de doctorat de l'université de Toulouse
Présentée par Julien OSMAN

Connaissances *a priori* pour l'exploitation d'images d'observation de la Terre à hautes résolutions spatiale, spectrale et temporelle

Jury

Jocelyn CHANUSSOT
Florence LE BER
Laurence MOY
Jean-Philippe GASTELLU-ETCHEGORRY
David SHEEREN
Jean-François DEJOUX
Jordi INGLADA

Rapporteur
Rapporteur
Rapporteur
Examineur
Examineur
Invité
Directeur de thèse

Sommaire

Introduction

Utilisation de l'information a priori

Les réseaux de logique de Markov

Expérimentations et résultats

Conclusions et perspectives

Sommaire

Introduction

Cartes d'occupation du sol

Utilisation des images satellite

Nouvelles données, nouveaux défis

Utilisation de l'information a priori

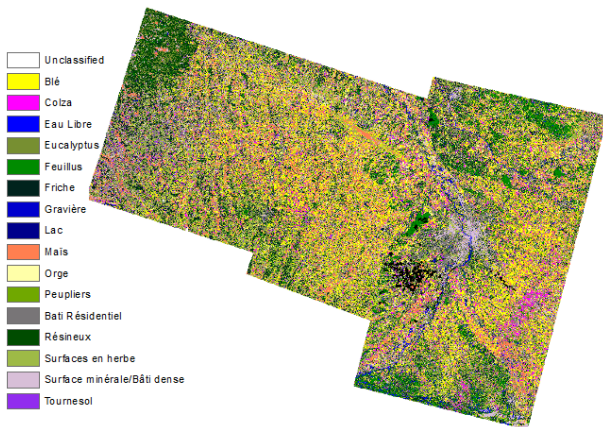
Les réseaux de logique de Markov

Expérimentations et résultats

Conclusions et perspectives

Cartes d'occupation du sol

"La couverture (bio-)physique de la surface des terres émergées."



Utilité

⇒ Première étape pour d'autres études :

▶ Agriculture

- ▶ Suivi des productions
- ▶ Surveillance des nappes phréatiques
- ▶ Surveillance de l'utilisation des nitrates
- ▶ Bilan des flux de carbone

▶ Surveillance des forêts

- ▶ Déforestation
- ▶ Incendies
- ▶ Maladies

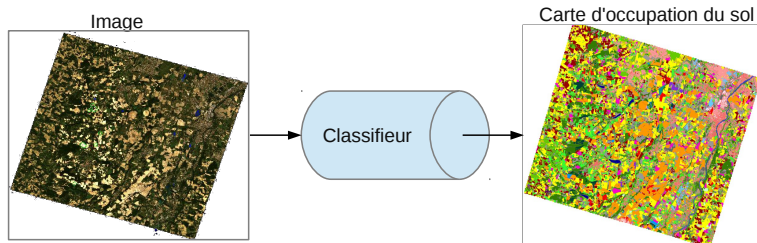
▶ Urbanisation

- ▶ Croissance urbaine
- ▶ Taux de surfaces artificialisées

▶ Gestion des risques

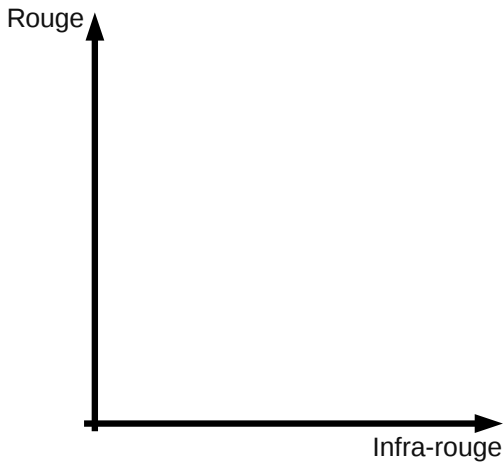
- ▶ Sécheresse
- ▶ Inondations

L'imagerie satellitaire

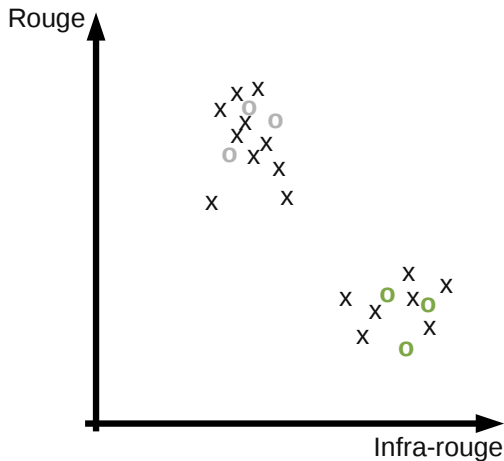


- ▶ Attribuer une classe à chaque individu
- ▶ Apprentissage : construction d'une fonction de décision
 - ▶ Supervisé : donnée de référence
 - ▶ Non-supervisé : étiquettes de classes non thématiques
- ▶ Inférence : attribution des classes

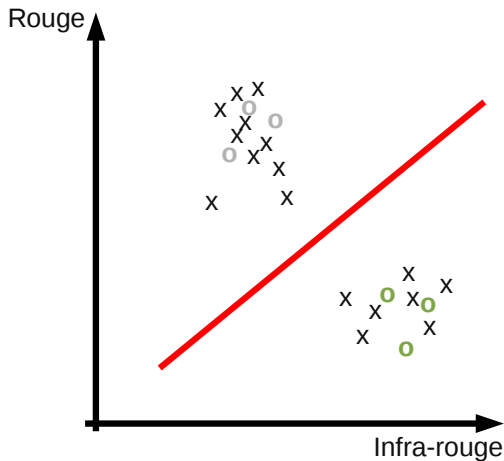
Fonctionnement des classifieurs



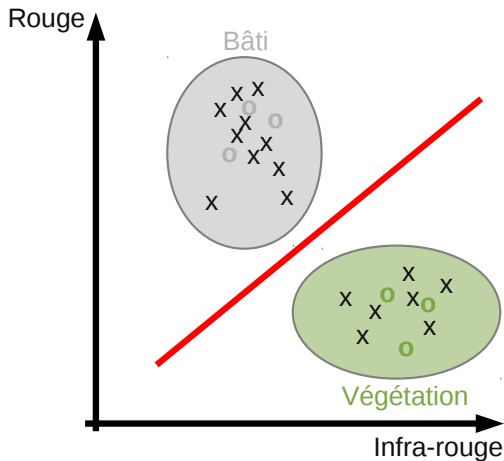
Fonctionnement des classifieurs



Fonctionnement des classifieurs



Fonctionnement des classifieurs



De nouvelles données

- ▶ Sentinel-2
 - ▶ Résolution spatiale : 10-60 m
 - ▶ Résolution spectrale : 13 bandes
 - ▶ Résolution temporelle : 5 jours (2 satellites à 10 jours)
 - ▶ Fauchée : 290 km

De nouvelles données

- ▶ Sentinel-2
 - ▶ Résolution spatiale : 10-60 m
 - ▶ Résolution spectrale : 13 bandes
 - ▶ Résolution temporelle : 5 jours (2 satellites à 10 jours)
 - ▶ Fauchée : 290 km
- ▶ Formosat-2
 - ▶ Résolution spatiale : 8 m
 - ▶ Résolution spectrale : 4 bandes
 - ▶ Résolution temporelle : 1 jour
 - ▶ Fauchée : 24 km

De nouveaux défis



⇒ De gros volumes de données

⇒ De grandes variétés de paysages

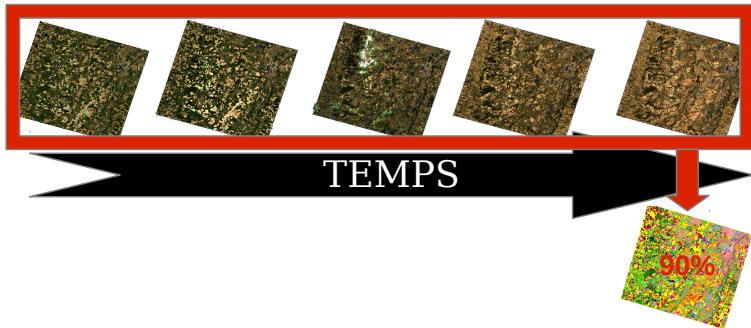
⇒ Donnée de référence inexistante

Objectifs de la thèse

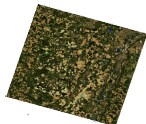
⇒ Exploiter ces nouvelles données de façon efficace pour la génération de cartes d'occupation du sol

- ▶ Concevoir et valider une approche
 - ▶ utilisable à grande échelle
 - ▶ traitant l'information rapidement
 - ▶ nécessitant peu de données de référence
- ▶ Estimer l'occupation du sol le plus tôt possible
 - ▶ classification au fil de l'eau
 - ▶ travailler avec peu d'images en début de saison

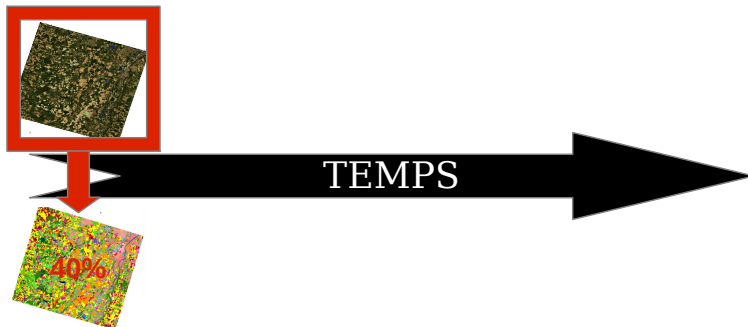
Classification au fil de l'eau



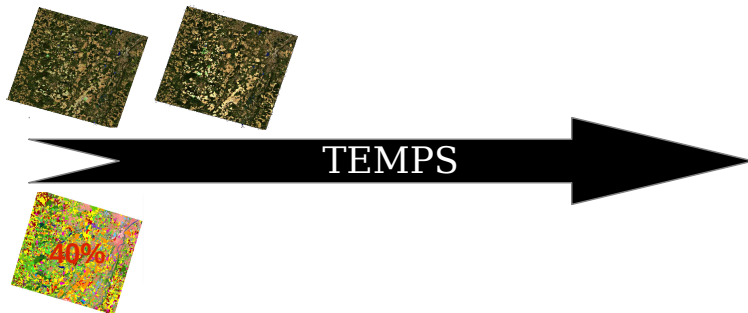
Classification au fil de l'eau



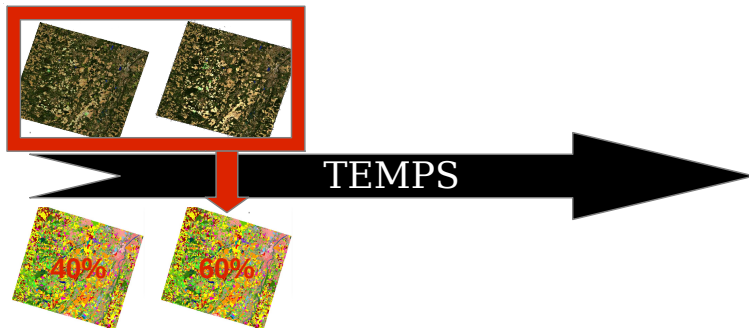
Classification au fil de l'eau



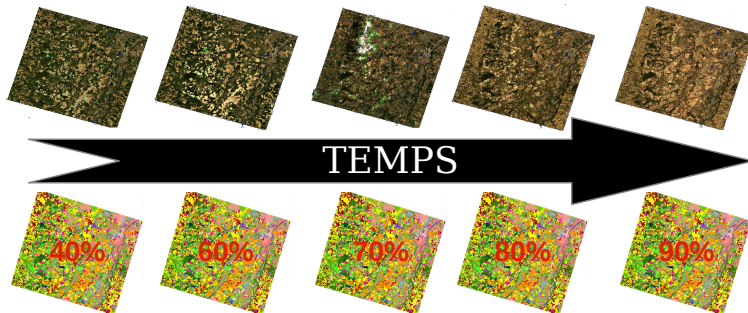
Classification au fil de l'eau



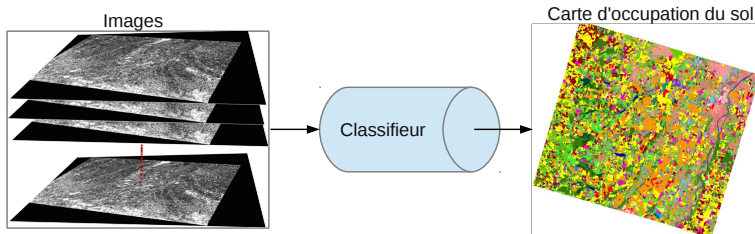
Classification au fil de l'eau



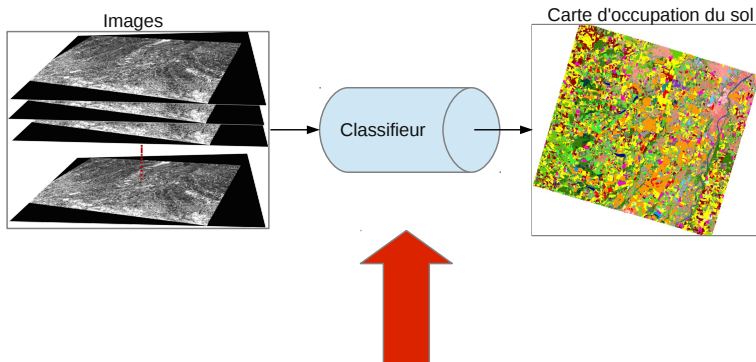
Classification au fil de l'eau



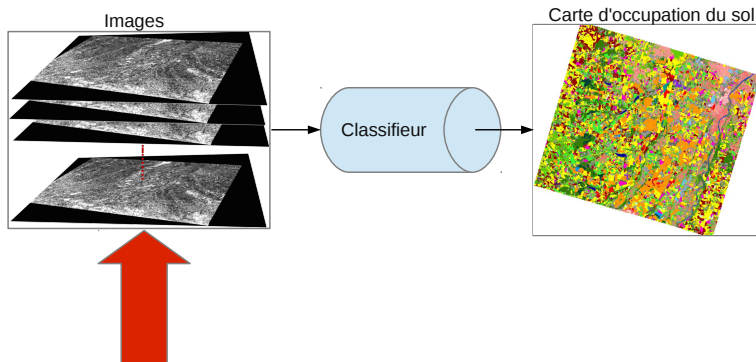
Deux approches pour améliorer la qualité des COS



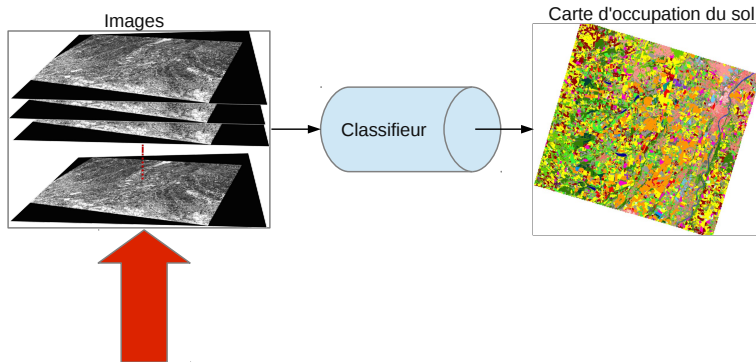
Deux approches pour améliorer la qualité des COS



Deux approches pour améliorer la qualité des COS

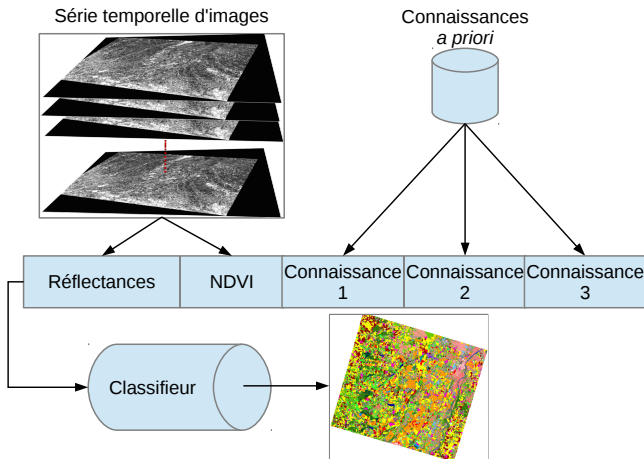


Deux approches pour améliorer la qualité des COS

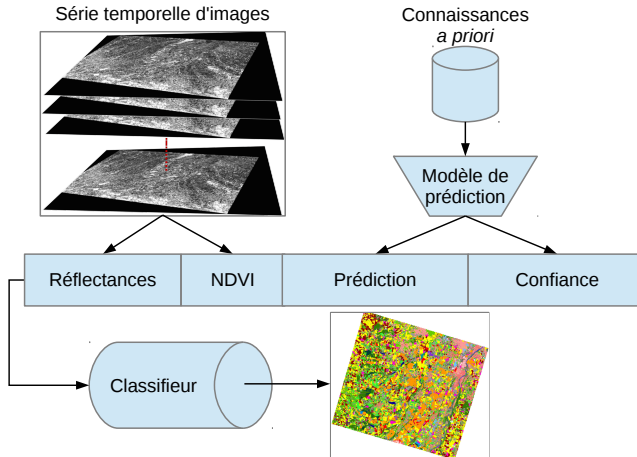


Étude restreinte au domaine de l'agriculture

Utilisation des connaissances



Utilisation des connaissances



Sommaire

Introduction

Utilisation de l'information a priori

Informations a priori

Les rotations de cultures

Cultures d'été/d'hiver

Comment exploiter ces connaissances ?

Les réseaux de logique de Markov

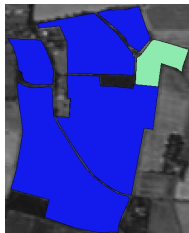
Expérimentations et résultats

Conclusions et perspectives

Informations a priori

- ▶ Culture au sol
 - ▶ Choix de l'exploitant
 - ▶ Les rotations de cultures
 - ▶ Les exploitations agricoles
 - ▶ L'irrigation
 - ▶ Les écorégions
 - ▶ Les itinéraires techniques
 - ▶ La topographie
 - ▶ Prix des graines
 - ▶ Rendements
 - ▶ Connaissances des télédéTECTEURS
 - ▶ La distinction cultures d'été/d'hiver
 - ▶ Forme des parcelles irriguées par pivot

Logique dans l'enchaînement des cultures



2006



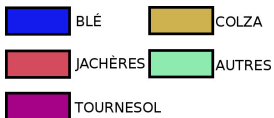
2007



2008



2009



Logique dans l'enchaînement des cultures



2006



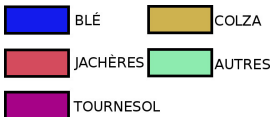
2007



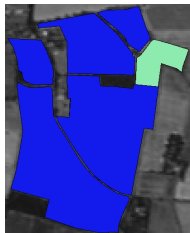
2008



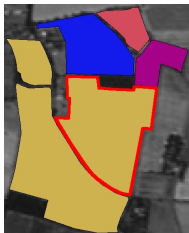
2009



Logique dans l'enchaînement des cultures



2006



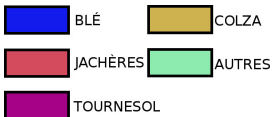
2007



2008



2009



Logique dans l'enchaînement des cultures



2006



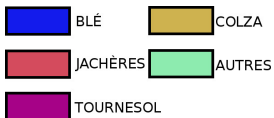
2007



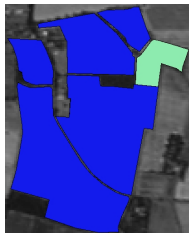
2008



2009



Logique dans l'enchaînement des cultures



2006



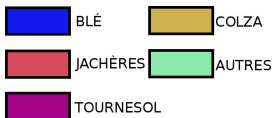
2007



2008



2009



Source de l'information

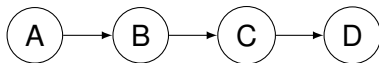
- ▶ Les bases de données graphiques comme le Registre Parcellaire Graphique (RPG)



- ▶ Les cartes d'occupation du sol des années précédentes

Modéliser les rotations de cultures

- ▶ Modèle statistique
 - ▶ Exploration de données
- ⇒ Modèles probabilistes graphiques (réseaux Bayésiens, chaînes de Markov, etc)



La distinction cultures d'été / cultures d'hiver

- ▶ Cultures d'hiver
 - ▶ Semées entre août et novembre
 - ▶ Passent l'hiver en ayant levé
 - ▶ Récoltées en été
 - ▶ Exemples : blé, orge, colza
 - ▶ Cultures d'été
 - ▶ Semées au printemps
 - ▶ Récoltées en automne
 - ▶ Exemples : maïs, tournesol
- ⇒ Signatures temporelles différentes
- ⇒ S'observe sur les images grâce au NDVI

Modéliser les cultures d'été/hiver

- ▶ Base de connaissances
 - ▶ Règles déterministes
- ⇒ Modèles logiques (calcul propositionnel, logique du 1^{er} ordre)

Comment exploiter ces connaissances ?

- ▶ Sources différentes
 - ▶ Bases de données
 - ▶ Images
 - ▶ Connaissances expertes
 - ▶ Approches de modélisation différentes
 - ▶ Statistique
 - ▶ Logique
- ⇒ Les réseaux de logique de Markov (MLN)

Sommaire

Introduction

Utilisation de l'information a priori

Les réseaux de logique de Markov

Présentation des MLN

Application aux rotations de cultures

Application aux règles sur les cultures d'été/hiver

Expérimentations et résultats

Conclusions et perspectives

Caractéristiques des MLN

- ▶ Ils combinent approches statistiques et approches logiques
- ▶ Leur mise à jour manuelle sans apprentissage est aisée
- ▶ Ils peuvent être fusionnés facilement

Exemple

- ▶ Règles :
 - ▶ Une culture irriguée est une culture d'été
 - ▶ Les cultures qui se succèdent ont les mêmes pratiques d'irrigation

Exemple

▶ Règles :

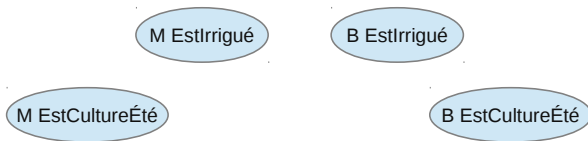
- ▶ $cultureAnneeN = x ; cultureAnneeM = y$
- ▶ $\forall x, x \text{ EstIrrigué} \Rightarrow x \text{ EstCultureÉté}$
- ▶ $\forall (x,y), x \text{ SuccèdeA } y \Rightarrow (x \text{ EstIrrigué} \Leftrightarrow y \text{ EstIrrigué})$

Exemple

- ▶ Règles :
 - ▶ $cultureAnneeN = x ; cultureAnneeM = y$
 - ▶ $1 ; \forall x, x \text{ EstIrrigué} \Rightarrow x \text{ EstCultureÉté}$
 - ▶ $3 ; \forall (x,y), x \text{ SuccèdeA } y \Rightarrow (x \text{ EstIrrigué} \Leftrightarrow y \text{ EstIrrigué})$

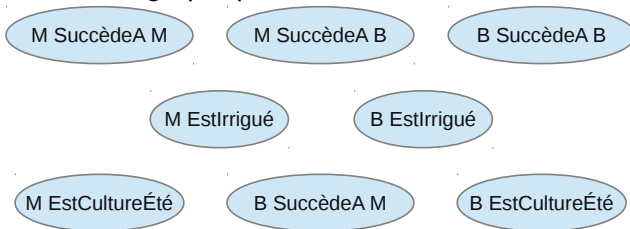
Exemple

- ▶ Règles :
 - ▶ $cultureAnneeN = x ; cultureAnneeM = y$
 - ▶ $1 ; \forall x, x \text{ EstIrrigué} \Rightarrow x \text{ EstCultureÉté}$
 - ▶ $3 ; \forall (x,y), x \text{ SuccèdeA } y \Rightarrow (x \text{ EstIrrigué} \Leftrightarrow y \text{ EstIrrigué})$
- ▶ 2 constantes : Maïs (M) et Blé (B)
- ▶ Représentation graphique



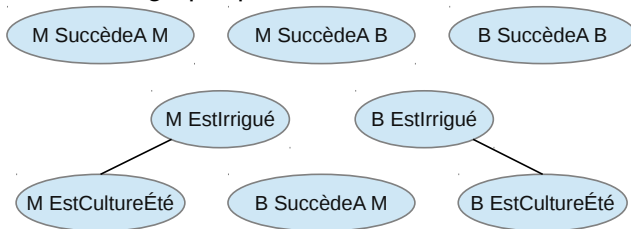
Exemple

- ▶ Règles :
 - ▶ $cultureAnneeN = x ; cultureAnneeM = y$
 - ▶ $1 ; \forall x, x \text{ EstIrrigué} \Rightarrow x \text{ EstCultureÉté}$
 - ▶ $3 ; \forall (x,y), x \text{ SuccèdeA } y \Rightarrow (x \text{ EstIrrigué} \Leftrightarrow y \text{ EstIrrigué})$
- ▶ 2 constantes : Maïs (M) et Blé (B)
- ▶ Représentation graphique



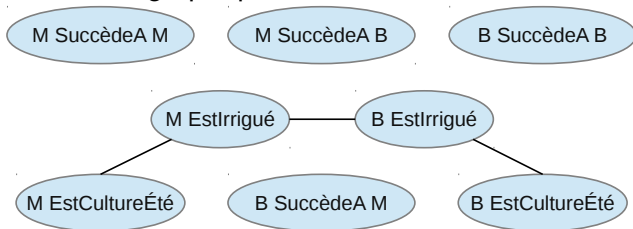
Exemple

- ▶ Règles :
 - ▶ $cultureAnneeN = x ; cultureAnneeM = y$
 - ▶ $1 ; \forall x, x \text{ EstIrrigué} \Rightarrow x \text{ EstCultureÉté}$
 - ▶ $3 ; \forall (x,y), x \text{ SuccèdeA } y \Rightarrow (x \text{ EstIrrigué} \Leftrightarrow y \text{ EstIrrigué})$
- ▶ 2 constantes : Maïs (M) et Blé (B)
- ▶ Représentation graphique



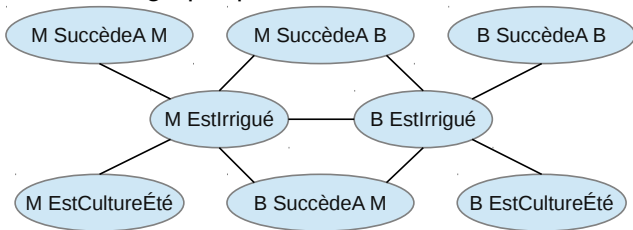
Exemple

- ▶ Règles :
 - ▶ $cultureAnneeN = x ; cultureAnneeM = y$
 - ▶ $1 ; \forall x, x \text{ EstIrrigué} \Rightarrow x \text{ EstCultureÉté}$
 - ▶ $3 ; \forall (x,y), x \text{ SuccèdeA } y \Rightarrow (x \text{ EstIrrigué} \Leftrightarrow y \text{ EstIrrigué})$
- ▶ 2 constantes : Maïs (M) et Blé (B)
- ▶ Représentation graphique



Exemple

- ▶ Règles :
 - ▶ $cultureAnneeN = x ; cultureAnneeM = y$
 - ▶ $1 ; \forall x, x \text{ EstIrrigué} \Rightarrow x \text{ EstCultureÉté}$
 - ▶ $3 ; \forall (x,y), x \text{ SuccèdeA } y \Rightarrow (x \text{ EstIrrigué} \Leftrightarrow y \text{ EstIrrigué})$
- ▶ 2 constantes : Maïs (M) et Blé (B)
- ▶ Représentation graphique



Définition

- ▶ Un MLN est un ensemble de paires (F, w) avec
 - ▶ F une formule de la logique du 1^{er} ordre
 - ▶ w un nombre réel
- ▶ Associé à un ensemble de constantes, il forme un champs aléatoire de Markov, avec
 - ▶ un nœud pour chaque formule atomique liée du MLN
 - ▶ une fonction caractéristique pour chaque formule liée du MLN avec le poids associé
- ▶ X suit alors la distribution de probabilité

$$\mathbf{P}(x) = \frac{1}{Z} \exp\left(\sum_i w_i n_i(x)\right)$$

- ▶ w_i : poids associé à la formule i
- ▶ n_i : Fonction caractéristique (f_i) définie comme étant le nombre de versions liées vraies de la formule i

Modéliser les rotations de cultures

- ▶ Une formule pour chaque combinaison possible de succession de cultures
- ▶ Un poids associé à chaque formule, symbolisant la probabilité d'apparition de la succession

$$(w_i : A_i \text{ et } B_i \text{ et } C_i \implies D_i)$$

- ▶ Les w_i sont appris à partir des données contenues dans le RPG ou des cartes d'occupation précédentes

Les règles expertes

Règles expertes sur les valeurs de NDVI entre avril et mai

$$\text{NDVI}_{t_i} \leq 0.3 \Rightarrow \text{culture d'été}$$

Les règles expertes

Règles expertes sur les valeurs de NDVI entre avril et mai

$$\text{NDVI}_{t_1} \leq 0.3 \Rightarrow \text{culture d'été}$$

$$\text{NDVI}_{t_1} \leq 0.3 \text{ ou } \text{NDVI}_{t_2} \leq 0.3 \Rightarrow \text{culture d'été}$$

Les règles expertes

Règles expertes sur les valeurs de NDVI entre avril et mai

$$\text{NDVI}_{t_1} \leq 0.3 \Rightarrow \text{culture d'été}$$

$$\text{NDVI}_{t_1} \leq 0.3 \text{ ou } \text{NDVI}_{t_2} \leq 0.3 \Rightarrow \text{culture d'été}$$

$$\left. \begin{array}{l} \text{NDVI}_{t_1} \leq 0.4 \text{ et } \text{NDVI}_{t_2} \leq 0.4 \text{ et } \text{NDVI}_{t_3} \leq 0.4 \\ \text{ou } (\text{NDVI}_{t_1} \leq 0.35 \text{ et } \text{NDVI}_{t_2} \leq 0.35) \\ \text{ou } (\text{NDVI}_{t_2} \leq 0.35 \text{ et } \text{NDVI}_{t_3} \leq 0.35) \\ \text{ou } (\text{NDVI}_{t_2} \leq 0.4 \text{ et } \text{NDVI}_{t_3} \leq 0.5) \\ \text{ou } \text{NDVI}_{t_1} \leq 0.3 \text{ ou } \text{NDVI}_{t_2} \leq 0.3 \text{ ou } \text{NDVI}_{t_3} \leq 0.3 \end{array} \right\} \Rightarrow \text{culture d'été}$$

MLN équivalent

- ▶ Une formule pour chaque règle experte, à laquelle est associé un poids
- ▶ Une formule pour définir la nature de chaque culture (été ou hiver) avec un poids élevé
 - ▶ Blé \implies culture d'été
 - ▶ Maïs \implies culture d'hiver

Sommaire

Introduction

Utilisation de l'information a priori

Les réseaux de logique de Markov

Expérimentations et résultats

Données

Expérimentations sur le MLN

Production des COS

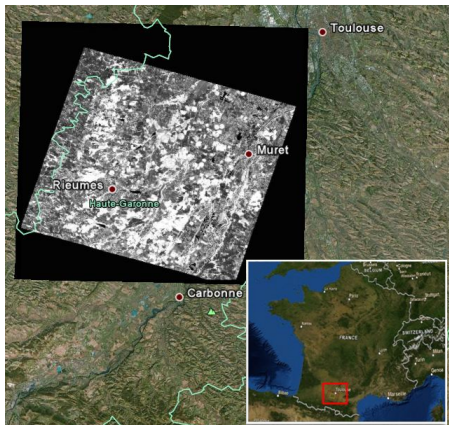
Conclusions et perspectives

Expérimentations

- ▶ Tester les caractéristiques des MLN
 - ▶ Combinent approches statistiques et approches logiques
 - ▶ Se mettent à jour sans apprentissage
 - ▶ Peuvent fusionner facilement
- ▶ Évaluer certains paramètres
 - ▶ Longueur optimale de rotation
 - ▶ Effets de la variété de paysages
 - ▶ Effets d'un important changement

Le site d'étude

Un site agricole au Sud-Ouest de la France

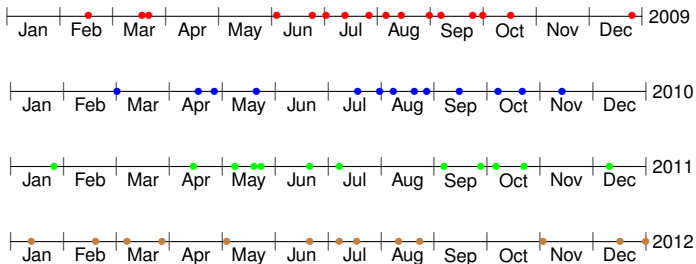


Cultures dominantes

- ▶ Blé
- ▶ Maïs
- ▶ Orge
- ▶ Colza
- ▶ Tournesol

Images satellite

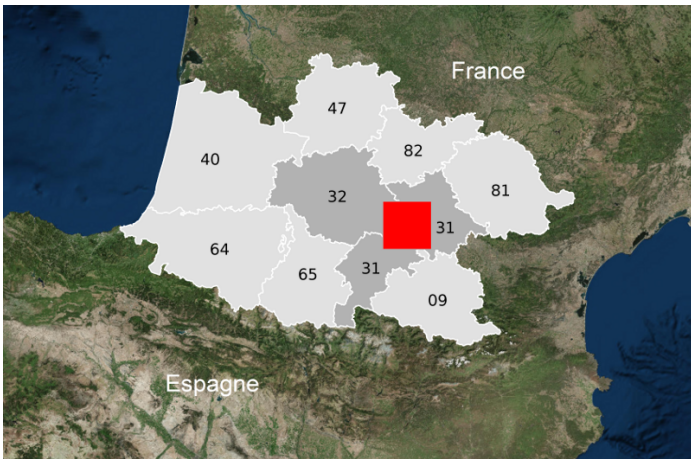
- ▶ Données Sentinel-2 pas encore disponibles
- ▶ Utilisation de données Formosat-2
- ▶ Données disponibles sur 4 années (2009 à 2012)



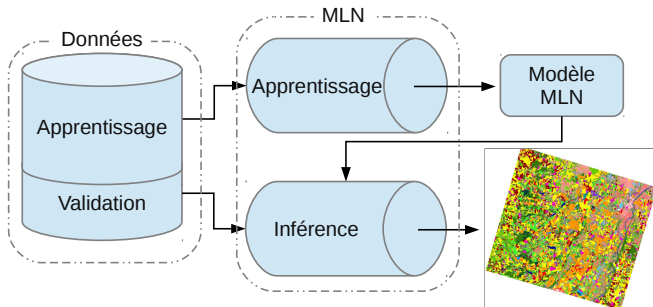
Le registre parcellaire graphique

- ▶ Apports
 - ▶ Contours géographiques des parcelles culturales
 - ▶ 28 classes agricoles
 - ▶ Disponible entre 2006 et 2012 sur tous les départements autour du site d'étude
- ▶ Inconvénients
 - ▶ Découpage en îlots. Pour travailler sur des parcelles, nous ignorons une grande partie des données
 - ▶ Ne couvre que les exploitants bénéficiant de la PAC
 - ▶ Peut contenir des erreurs (positionnement des îlots, classes mal renseignées)

Zones d'apprentissage

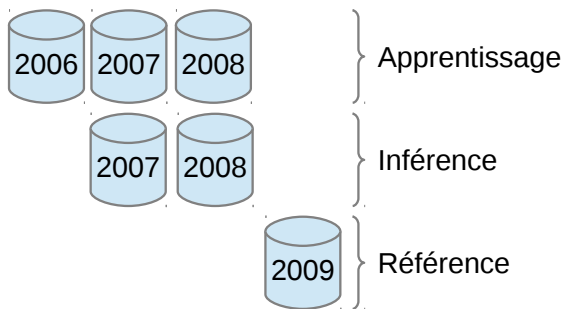


Configuration expérimentale



Utilisation de l'indice kappa (κ) pour l'évaluation des résultats

Utilisation du RPG



- ▶ Apprentissage avec 4000 parcelles de la zone moyenne
- ▶ Inférence avec 4000 autres parcelles de la zone moyenne
- ▶ Étude des 5 cultures dominantes (blé, maïs, orge, colza, tournesol)

Longueur optimale de rotation

- ▶ Rotations sur 3 années : précision = 64%
 - ▶ Rotations sur 4 années : précision = 59%
 - ▶ Rotations sur 5 années : précision = 58%
- ⇒ Travailler sur de longues rotations favorise les cultures qui sont dans des rotations complexes

-	Bl	Ma	Or	Co	To
Bl	72	10	6	4	8
Ma	4	85	3	3	5
Or	31	13	10	12	34
Co	7	8	10	27	49
To	12	8	6	11	62

Rotations sur 3 années

-	Bl	Ma	Or	Co	To
Bl	59	12	9	9	11
Ma	11	71	6	6	7
Or	23	18	22	14	22
Co	8	8	11	30	43
To	8	9	10	9	65

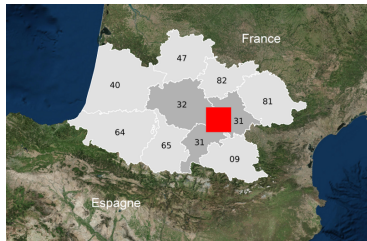
Rotations sur 4 années

-	Bl	Ma	Or	Co	To
Bl	60	11	9	8	12
Ma	12	69	6	7	6
Or	19	14	26	12	29
Co	12	13	11	23	42
To	12	12	6	9	61

Rotations sur 5 années

Effet de la variété du paysage

- ▶ Apprentissage sur les 3 zones : tailles et variétés de paysages différentes
- ▶ Inférence sur la zone Formosat
- ▶ Rotations sur 3, 4 et 5 années
- ▶ 27 combinaisons

Valeurs de κ

Zone Formosat					Zone moyenne				
-	2009	2010	2011	2012	-	2009	2010	2011	2012
3 années	0.51	0.58	0.54	0.60	3 années	0.53	0.57	0.51	0.58
4 années	-	0.57	0.53	0.61	4 années	-	0.57	0.52	0.59
5 années	-	-	0.54	0.55	5 années	-	-	0.51	0.54

Zone élargie				
-	2009	2010	2011	2012
3 années	0.50	0.56	0.52	0.58
4 années	-	0.50	0.46	0.53
5 années	-	-	0.43	0.43

Mise à jour du MLN sans apprentissage

- ▶ Exemple d'un changement dans les pratiques agricoles : simuler une loi interdisant l'enchaînement de 4 années de maïs (rotation très présente dans la zone d'étude)
- ▶ Pas de donnée d'apprentissage, modélisable par expertise
- ▶ Forte réduction du poids associé à la règle

Maïs et Maïs et Maïs \implies Maïs

- ▶ Influence sur les résultats de l'inférence

-	Original	Modifié
Maïs	0.60	0.0014
Blé	0.11	0.28
Tournesol	0.11	0.28
Colza	0.088	0.22
Orge	0.089	0.23

Probabilité d'apparition de chaque culture suite à 3 années de maïs.

- ▶ Les autres rotations ne sont pas influencées par le changement

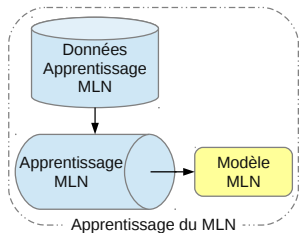
Fusion des modèles

- ▶ Construction du modèle sur les cultures d'été
- ▶ Apprentissage du modèle sur les rotations de cultures dans les mêmes conditions
- ▶ Fusion des deux modèles
- ▶ Inférence des 3 modèles avec les mêmes données de validation

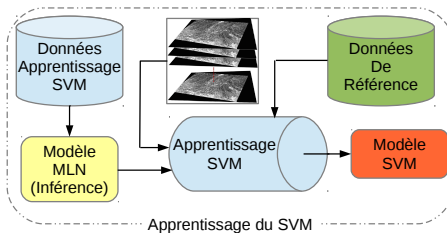
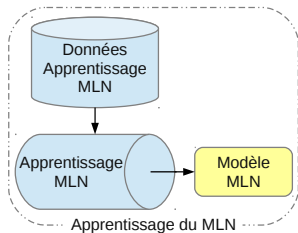
Valeurs de κ

version méthode	1 date	2 dates	3 dates
Règles expertes	0.16	0.15	0.14
Rotations de cultures	0.45		
Fusion règles et rotations	0.51	0.53	0.56

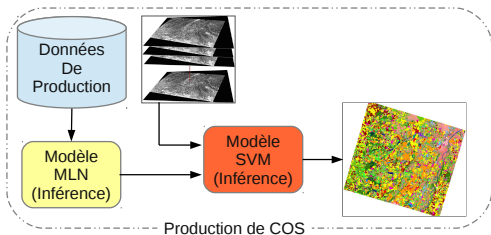
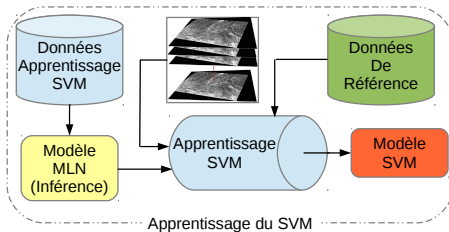
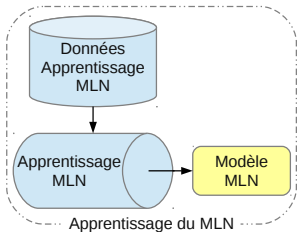
Configuration expérimentale



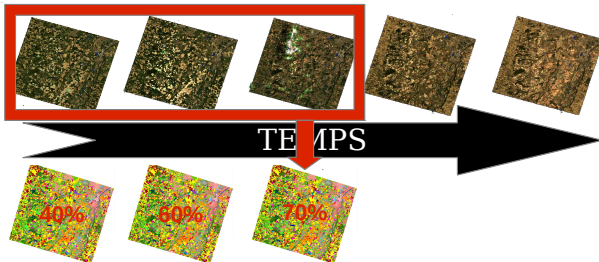
Configuration expérimentale



Configuration expérimentale

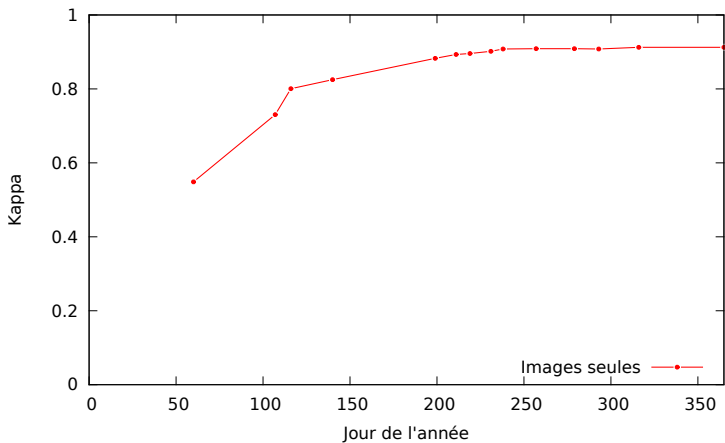


Fil de l'eau simulé

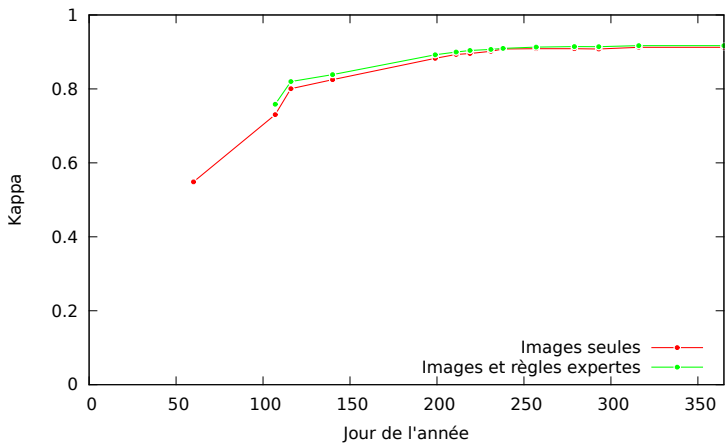


- ▶ Simulation d'une classification au fil de l'eau
 - ▶ Réalisation de plusieurs classifications
 - ▶ Une nouvelle COS à l'arrivée de chaque image

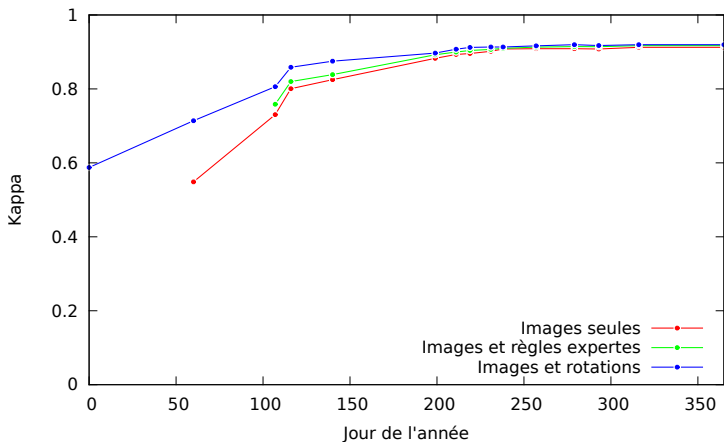
Résultats



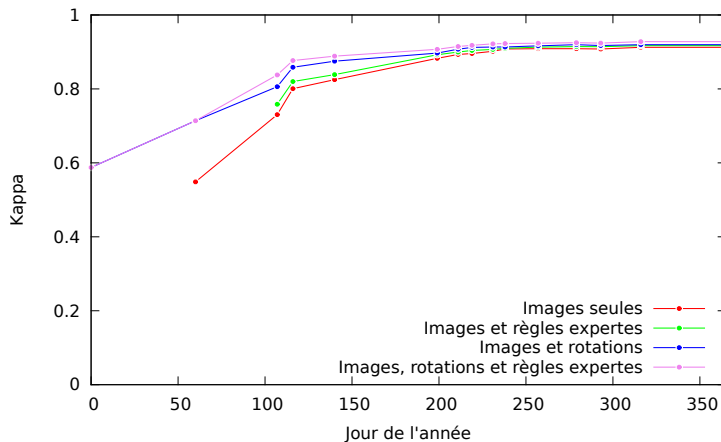
Résultats



Résultats



Résultats



Sommaire

Introduction

Utilisation de l'information a priori

Les réseaux de logique de Markov

Expérimentations et résultats

Conclusions et perspectives

Conclusions

Perspectives

Conclusions

- ▶ Identification des connaissances *a priori* ajoutant de l'information au procédé de classification
 - ▶ Les caractériser
 - ▶ Les modéliser
 - ▶ Exploiter les modèles
- ▶ Choix d'un outil adapté à la modélisation et à l'exploitation de ces connaissances
 - ▶ Combiner approches logiques et approches statistiques
 - ▶ Fusion et mise à jour des modèles
 - ▶ Classification non supervisée sans reconnaissance des classes *a posteriori*
- ▶ Validation de l'approche
 - ▶ Mise en place d'un banc d'essai rigoureux
 - ▶ Analyse et discussion détaillées des résultats

Perspectives

- ▶ Pour le domaine de l'agriculture
 - ▶ Introduire d'autres connaissances
 - ▶ L'irrigation
 - ▶ Les écorégions
 - ▶ Les itinéraires techniques
- ▶ Pour les autres applications
 - ▶ Suivi des forêts tropicales
 - ▶ Étalement urbain
 - ▶ L'occupation du sol en général
- ▶ Pour la méthode
 - ▶ Utiliser moins de données de référence
 - ▶ Apprentissage de la structure du réseau
 - ▶ Construction dynamique du MLN
 - ▶ Optimisation des paramètres et primitives du classifieur
 - ▶ Prendre en compte la configuration spatiale

Merci pour votre attention



Annexes

Introduction

Imagerie satellitaire

Les MLN

Les MRF

Logique 1^{er} ordre

Intuition sur les MLN

Application aux rotations de cultures

Résultats

Temps de calcul

Matrices de confusion

Sommaire

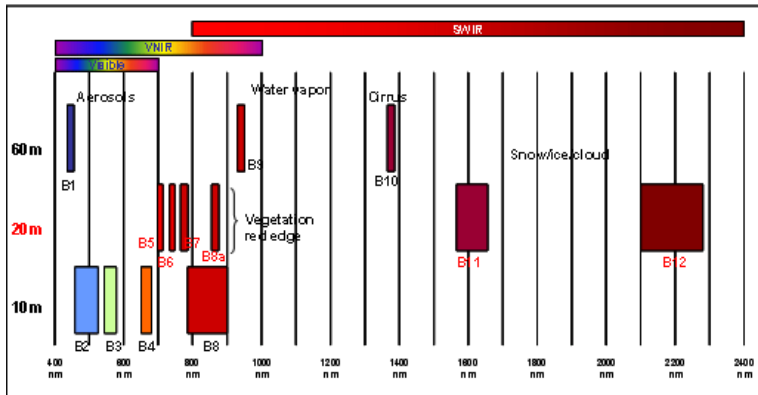
Introduction

Imagerie satellitaire

Les MLN

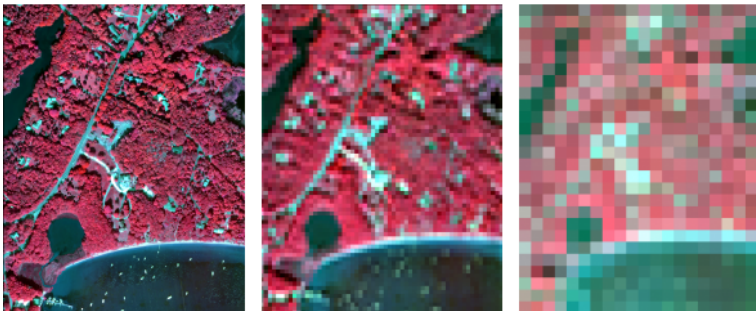
Résultats

Sentinel-2 : Bandes spectrales



Informations contenues dans les images

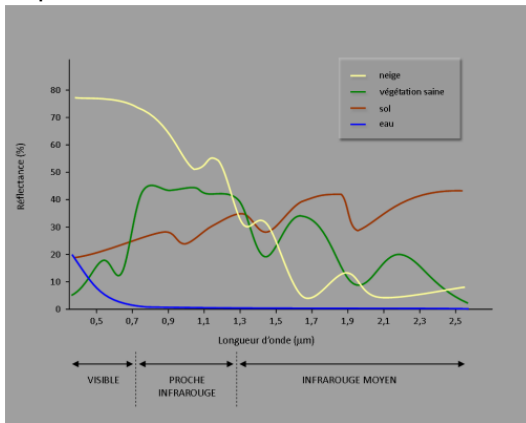
► Signature spatiale



Exemple de résolutions spatiales : 1 m, 10 m et 30 m

Informations contenues dans les images

- ▶ Signature spatiale
- ▶ Signature spectrale



Informations contenues dans les images

- ▶ Signature spatiale
- ▶ Signature spectrale
- ▶ Signature temporelle

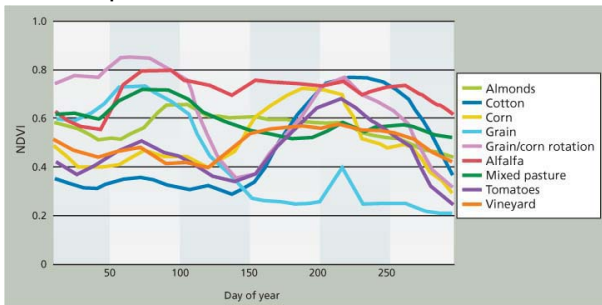


Fig. 2. Normalized difference vegetation index (NDVI) values by day of year for all time periods, for nine agricultural land-cover classes.

Sommaire

Introduction

Les MLN

Les MRF

Logique 1^{er} ordre

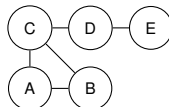
Intuition sur les MLN

Application aux rotations de cultures

Résultats

Les champs aléatoires de Markov (MRF)

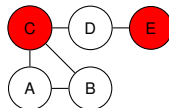
- ▶ Soit $G = (N, A)$ un graphe non orienté



- ▶ Soit $X = (X_i)_{i \in N}$ un ensemble de variables aléatoires (une pour chaque sommet de G), respectant la propriété de Markov, c'est à dire une de ces 3 règles :

Les champs aléatoires de Markov (MRF)

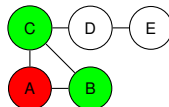
- ▶ Soit $G = (N, A)$ un graphe non orienté



- ▶ Soit $X = (X_i)_{i \in N}$ un ensemble de variables aléatoires (une pour chaque sommet de G), respectant la propriété de Markov, c'est à dire une de ces 3 règles :
 - ▶ Deux variables dont les sommets associés ne sont pas reliées par un arc sont indépendantes conditionnellement à toutes les autres variables

Les champs aléatoires de Markov (MRF)

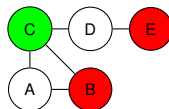
- ▶ Soit $G = (N, A)$ un graphe non orienté



- ▶ Soit $X = (X_i)_{i \in N}$ un ensemble de variables aléatoires (une pour chaque sommet de G), respectant la propriété de Markov, c'est à dire une de ces 3 règles :
 - ▶ Deux variables dont les sommets associés ne sont pas reliés par un arc sont indépendantes conditionnellement à toutes les autres variables
 - ▶ Une variable est indépendante de toutes les autres conditionnellement à son voisinage

Les champs aléatoires de Markov (MRF)

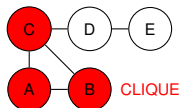
- ▶ Soit $G = (N, A)$ un graphe non orienté



- ▶ Soit $X = (X_i)_{i \in N}$ un ensemble de variables aléatoires (une pour chaque sommet de G), respectant la propriété de Markov, c'est à dire une de ces 3 règles :
 - ▶ Deux variables dont les sommets associés ne sont pas reliés par un arc sont indépendantes conditionnellement à toutes les autres variables
 - ▶ Une variable est indépendante de toutes les autres conditionnellement à son voisinage
 - ▶ Deux variables sont indépendantes conditionnellement à S si tout chemin entre ces deux variables passe par S

Les champs aléatoires de Markov (MRF)

- ▶ Soit $G = (N, A)$ un graphe non orienté



- ▶ Soit $X = (X_i)_{i \in N}$ un ensemble de variables aléatoires (une pour chaque sommet de G), respectant la propriété de Markov, c'est à dire une de ces 3 règles :
 - ▶ Deux variables dont les sommets associés ne sont pas reliés par un arc sont indépendantes conditionnellement à toutes les autres variables
 - ▶ Une variable est indépendante de toutes les autres conditionnellement à son voisinage
 - ▶ Deux variables sont indépendantes conditionnellement à S si tout chemin entre ces deux variables passe par S

Théorème de Hammersley-Clifford

- ▶ Une distribution de probabilités dont la densité est strictement positive satisfait une des propriétés de Markov si et seulement si sa densité est une mesure de Gibbs
- ▶ Elle peut alors s'écrire sous une forme factorisée sur les cliques de G :

$$P(X = x) = \frac{1}{Z} \exp\left(\sum_i w_i f_i(x)\right)$$

- ▶ f_i une fonction caractéristique dépendant de la configuration des cliques de G
- ▶ w_i un poids lié au voisinage du nœud associé à la variable aléatoire X_i
- ▶ Z une fonction de partition (constante de normalisation)

La logique du 1^{er} ordre

- ▶ Permet de représenter des formules logiques à l'aide de symboles :
 - ▶ Les variables (*cultureAnneeN*)
 - ▶ Les constantes : valeurs que les variables peuvent prendre (Maïs)
 - ▶ Les prédicats : propriété d'une ou plusieurs variables (**EstIrrigué**)
- ▶ Quelques définitions :
 - ▶ Un atome est une formule qui n'est composée que d'un seul prédicat (formule atomique)
 - ▶ Un énoncé est une formule dans laquelle les variables ont été liées à une constante
 - ▶ On parle de domaine lorsqu'à chaque énoncé est associé un état vrai ou faux

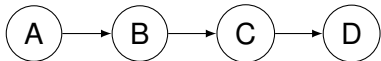
Unir les deux approches

- ▶ Une base de connaissances logiques est un ensemble de règles strictes.
 - ▶ Rendons ces règles souples : si un domaine ne respecte pas une règle, il devient moins probable, et non impossible.
 - ▶ Attribuons à chaque règle un poids : plus le poids est grand, plus la contrainte est forte.
- ⇒ $\mathbf{W}(\text{domaine}) \propto \exp(\sum \text{poids des règles satisfaites})$

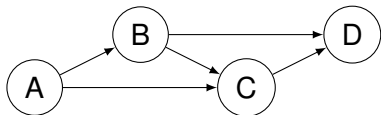
Unir les deux approches

- ▶ Une base de connaissances logiques est un ensemble de règles strictes.
 - ▶ Rendons ces règles souples : si un domaine ne respecte pas une règle, il devient moins probable, et non impossible.
 - ▶ Attribuons à chaque règle un poids : plus le poids est grand, plus la contrainte est forte.
- ⇒ $W(\text{domaine}) \propto \exp(\sum \text{poids des règles satisfaites})$
- ▶ Exemples (les poids des règles valent 1) :
 - ▶ Domaine : Maïs **EstIrrigué** ; Blé **EstIrrigué**. $W \propto e^1$
 - ▶ Domaine : Maïs **EstIrrigué** ; Blé **NonIrrigué**. $W \propto e^2$

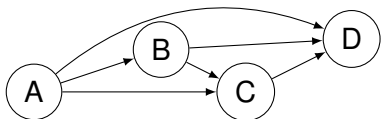
Modèles possibles pour les rotations de cultures



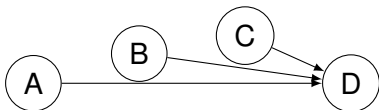
(e) Chaîne de Markov du 1^{er} ordre



(f) Chaîne de Markov du 2^{ème} ordre



(g) Chaîne de Markov du 3^{ème} ordre



(h) Effet commun

Sommaire

Introduction

Les MLN

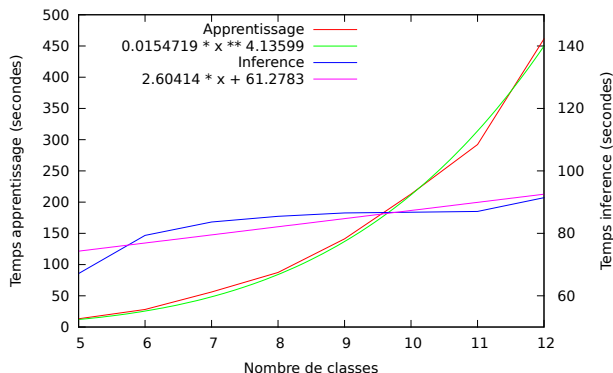
Résultats

Temps de calcul

Matrices de confusion

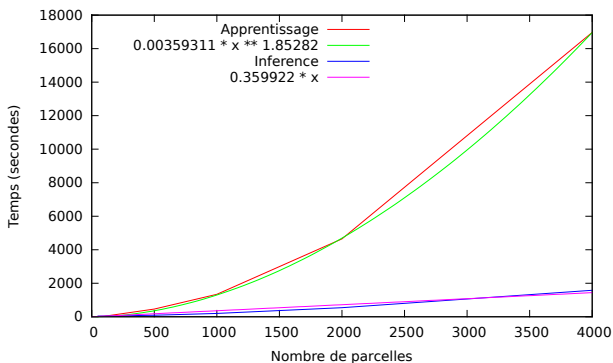
Temps d'apprentissage et d'inférence

- ▶ Série d'apprentissages et d'inférences
- ▶ Influence du nombre de classes



Temps d'apprentissage et d'inférence

- ▶ Série d'apprentissages et d'inférences
- ▶ Influence du nombre de parcelles



Matrices de confusion

- ▶ Matrice de confusion obtenue le 1^{er} janvier 2010 en utilisant uniquement la connaissance sur les rotations de cultures

-	Ble	Maïs	Orge	Colza	Trnsol
Ble	69	1	24	6	1
Maïs	5	36	44	2	12
Orge	11	0	72	12	4
Colza	6	0	22	68	4
Trnsol	6	1	24	19	50
-	Précision	59.20 %	Kappa	0.5876	-

- ▶ Matrice de confusion obtenue le 2 mars 2010 en utilisant uniquement la première image de l'année

-	Ble	Maïs	Orge	Colza	Trnsol
Ble	49	2	26	14	9
Maïs	1	73	5	3	18
Orge	16	5	61	5	13
Colza	18	3	17	52	11
Trnsol	11	36	7	5	42
-	Précision	55.35 %	Kappa	0.5487	-

Matrices de confusion

- ▶ Matrice de confusion obtenue le 2 mars 2010 en utilisant uniquement la première image de l'année

-	Ble	Maïs	Orge	Colza	Trnsol
Ble	49	2	26	14	9
Maïs	1	73	5	3	18
Orge	16	5	61	5	13
Colza	18	3	17	52	11
Trnsol	11	36	7	5	42
-	Précision	55.35 %	Kappa	0.5487	-

- ▶ Matrice de confusion obtenue le 2 mars 2010 en utilisant la première image de l'année et la connaissance sur les rotations de cultures

-	Ble	Maïs	Orge	Colza	Trnsol
Ble	66	2	22	5	4
Maïs	3	84	5	1	8
Orge	11	2	70	10	6
Colza	7	2	14	73	4
Trnsol	7	9	11	7	66
-	Précision	71.92 %	Kappa	0.714	-

Matrices de confusion

- ▶ Matrice de confusion obtenue le 31 décembre 2010 en utilisant uniquement la série d'images

-	Ble	Maïs	Orge	Colza	Trnsol
Ble	86	0	11	2	1
Maïs	1	96	1	0	2
Orge	7	0	90	3	1
Colza	2	0	5	93	0
Trnsol	2	1	2	1	94
-	Précision	91.56 %	Kappa	0.9125	-

- ▶ Matrice de confusion obtenue le 31 décembre 2010 en utilisant les images et les connaissances *a priori*

-	Ble	Maïs	Orge	Colza	Trnsol
Ble	89	0	7	2	2
Maïs	1	97	1	0	2
Orge	6	0	90	2	2
Colza	2	0	3	94	1
Trnsol	2	1	1	1	95
-	Précision	93.06 %	Kappa	0.9279	-