

# PRÉDICTION DES SERVICES ÉCOSYSTÉMIQUES DANS LES PAYSAGES AGRICOLES PAR TÉLÉDÉTECTION HYPERSPECTRALE

6<sup>e</sup> COLLOQUE GROUPE HYPERSPECTRAL SFPT-GH

---

M. Fauvel, R. Duflot, T. Uezato, N. Dobigeon, A. Vialatte, D. Sheeren, M. San Cristobal et M. Goulard

CONTEXTE

## CONTEXTE

CARTOGRAPHIE DES SERVICES ECOSYSTÉMIQUES PAR TÉLÉDÉTECTION

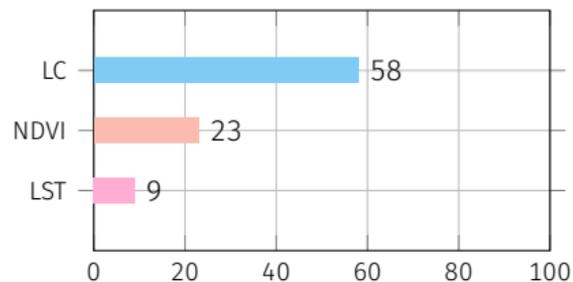
## Cartographie des services écosystémiques (SE)

- Surveillance de l'état et contrôle de l'évolution des SE.
- Nécessaire pour l'évaluation et l'élaboration des politiques d'aménagement des territoires :
  - ▶ Échelles nationales et internationales,
  - ▶ Méthodes comparables.

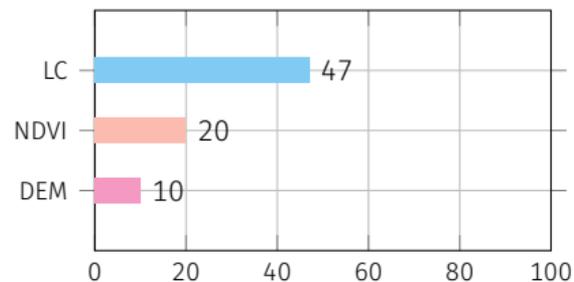
## Méthodes de cartographie

- Relevés « terrain » :
  - ▶ Précis,
  - ▶ Faible couverture spatiale et répétitivité temporelle,
  - ▶ Comparaison difficile.
- Données d'observation de la terre :
  - ▶ Couverture spatiale importante et haute répétitivité temporelle,
  - ▶ Coût financier faible.

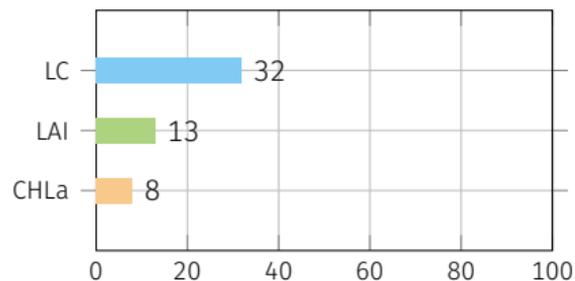
## Services de provisions



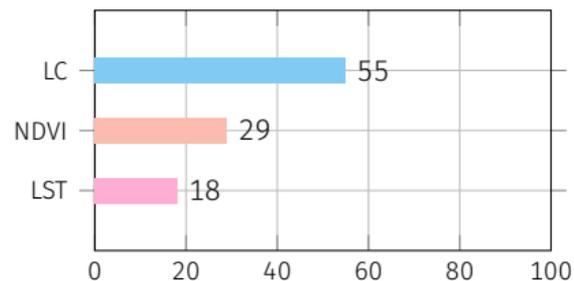
## Services de régulations

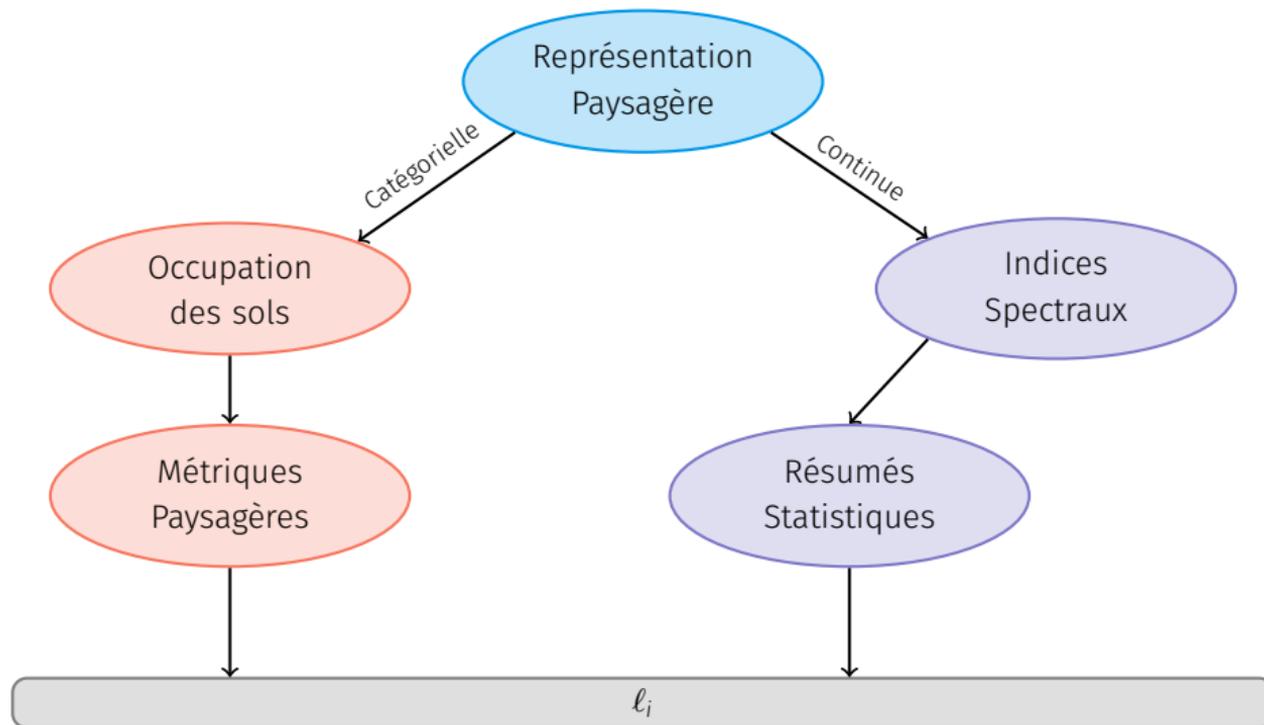


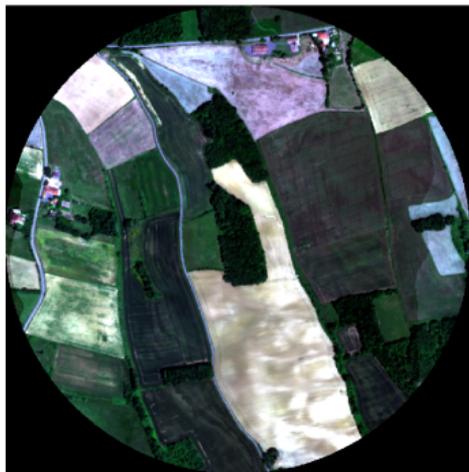
## Services de soutiens



## Services culturels







CONTEXTE

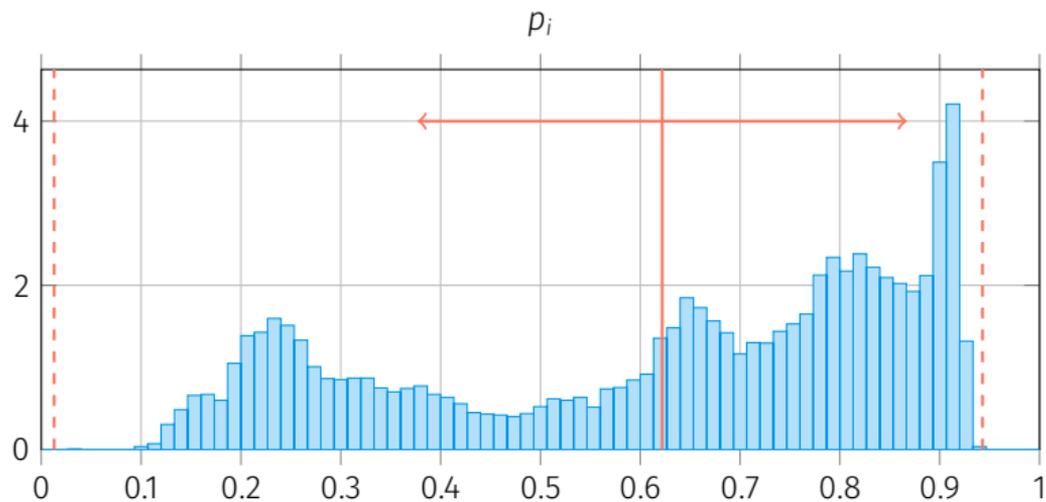
LIMITE DES APPROCHES ACTUELLES

### Approche Catégorielle

- ✓ Composition et configuration
- × Cartographie d'OCS manuelle
- × Pas de distinction pour un même OCS

### Approche Continue

- ✓ Pas d'OCS
- ✓ Etat et structure de la végétation
- × Limitée à la composition

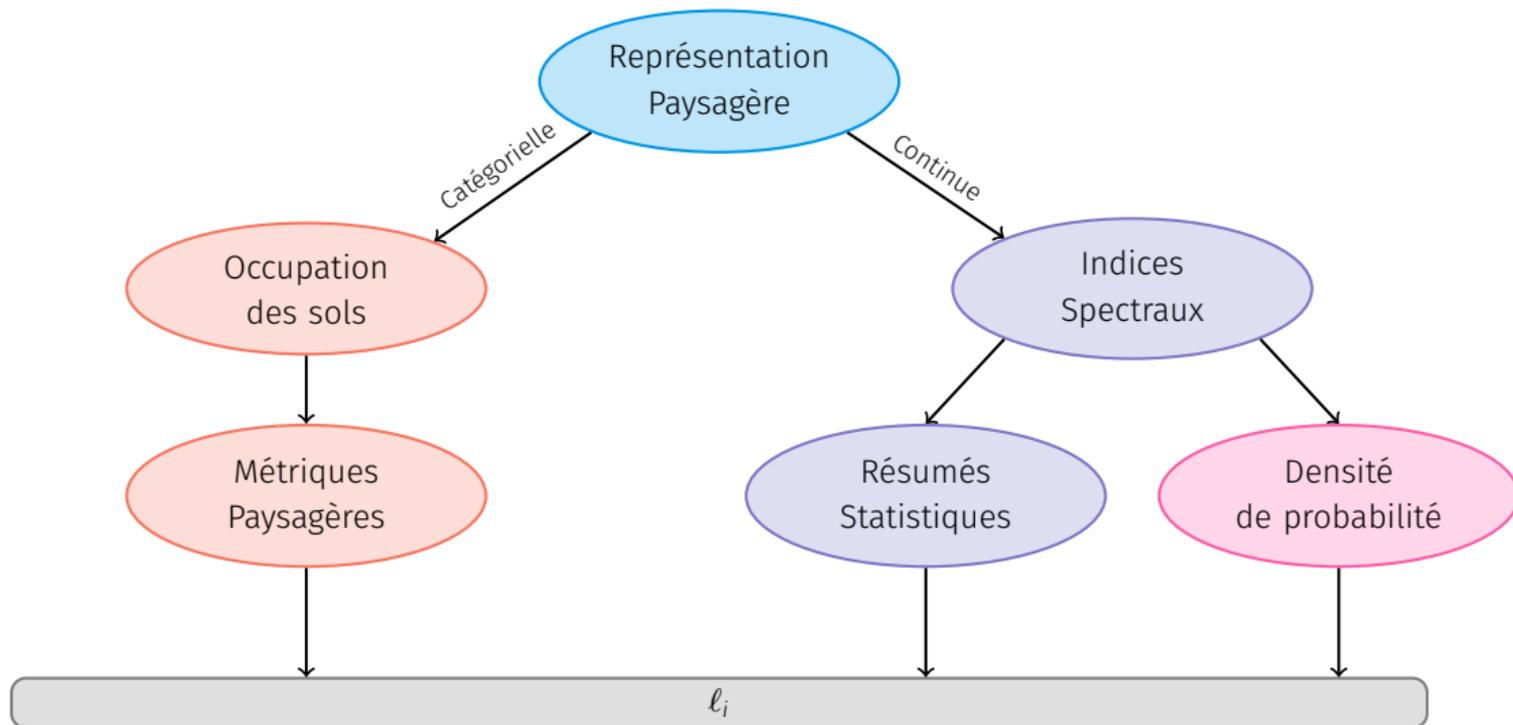


- Discrétisation de la distribution des valeurs des pixels du paysage  $\ell_i$

$$\ell_i = [\mu, \sigma, \max, \min]$$

### Modélisation continue de la distribution

1. Modélisation de la distribution par lois de mélange Gaussiennes (GMM)
2. Mesure de distance entre GMM
3. Algorithme de prédiction non-linéaire



**MODÉLISATION CONTINUE**

**MODÉLISATION CONTINUE**  
**MODÉLISATION DE LA RADIOMÉTRIE**

- On suppose qu'un pixel est distribué selon une loi de mélange (C le nombre de clusters)

$$p(\mathbf{x}) = \sum_{c=1}^C p(\mathbf{x}|c)p(c)$$

- On en déduit la distribution de  $\mathbf{x}$  pour le paysage  $\ell$  :

$$p(\mathbf{x}, \ell) \propto p(\mathbf{x}|\ell) = \sum_{c=1}^C p(\mathbf{x}|c, \ell)p(c|\ell) = \sum_{c=1}^C p(\mathbf{x}|c)p(c|\ell)$$

*On suppose que*

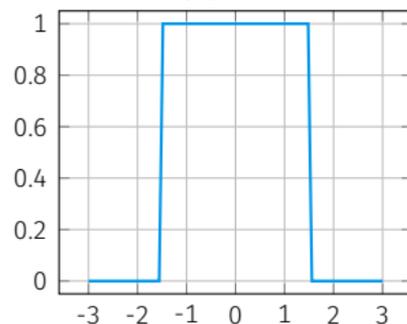
- ▶ Les paysages sont équiprobables,
  - ▶ La loi conditionnelle au cluster ne dépend pas du paysage.
- On peut estimer  $p(c|\ell)$  par :

$$p(c|\ell) = \frac{\sum_{i=1}^n p(c|\mathbf{x}_i)\Psi_{\ell}(\mathbf{x}_i)}{\sum_{i=1}^n \Psi_{\ell}(\mathbf{x}_i)}$$

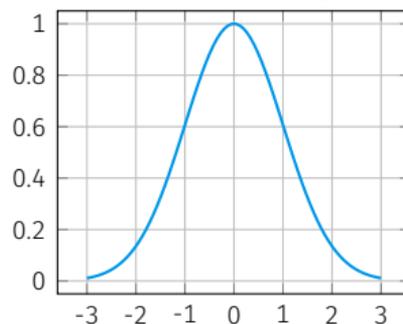
où  $\Psi_{\ell}$  est une fonction d'appartenance au paysage  $\ell$ .

Fonction de la distance géographique  $d_i$  du pixel  $x_i$  au centre du paysage  $\ell$

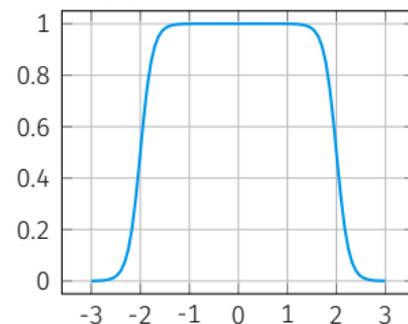
$$\Psi_{\ell}(x_i) = 1 \vee 0$$



$$\Psi_{\ell}(x_i) = \exp\left(-\frac{d_i^2}{\sigma^2}\right)$$



$$\Psi_{\ell}(x_i) = \text{Double logistique}$$



**MODÉLISATION CONTINUE**

**MODÉLISATION DU PAYSAGE**

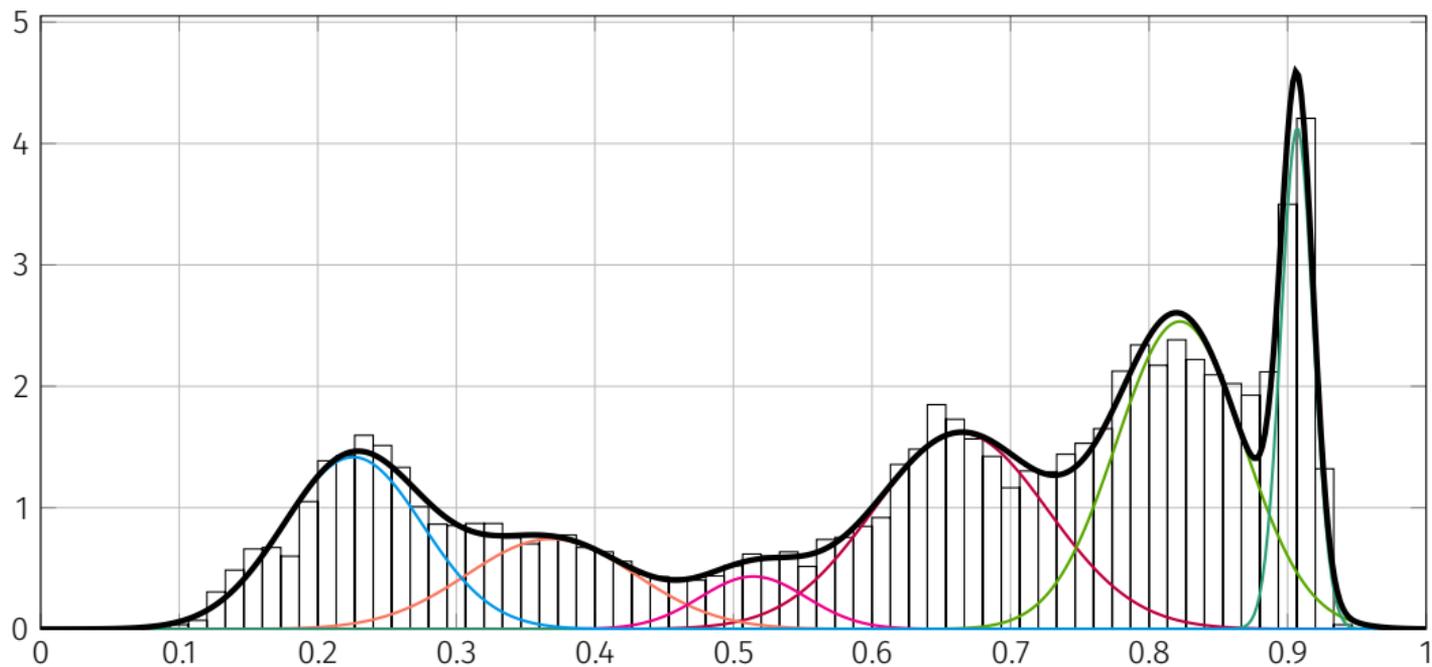
- Un paysage  $\ell$  est modélisé par  $p(\mathbf{x}|\ell)$ . On suppose de plus que les lois conditionnelles aux clusters suivent une loi normale multivariée :  $p(\mathbf{x}|c) \sim \mathcal{N}(\boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c)$ .
- Un produit scalaire entre deux paysages  $\ell$  et  $\ell'$  peut se définir alors comme :

$$\begin{aligned}
 \langle \ell, \ell' \rangle &= \int_{\mathbb{R}^d} p(\mathbf{x}|\ell)p(\mathbf{x}|\ell')d\mathbf{x} \\
 &= \sum_{c,c'=1}^{C,C} p(c|\ell)p(c'|\ell') \frac{\exp \left[ -0.5(\boldsymbol{\mu}_c - \boldsymbol{\mu}_{c'})^\top (\boldsymbol{\Sigma}_c + \boldsymbol{\Sigma}_{c'})^{-1} (\boldsymbol{\mu}_c - \boldsymbol{\mu}_{c'}) \right]}{|2\pi(\boldsymbol{\Sigma}_c + \boldsymbol{\Sigma}_{c'})|^{0.5}} \\
 &= \sum_{c,c'=1}^{C,C} p(c|\ell)p(c'|\ell')D_{cc'} \\
 &= \mathbf{p}_\ell^\top \mathbf{D} \mathbf{p}_{\ell'} \text{ avec } \mathbf{p}_\ell \in \mathbb{R}^C \text{ et } [\mathbf{p}_\ell]_c = p(c|\ell).
 \end{aligned}$$

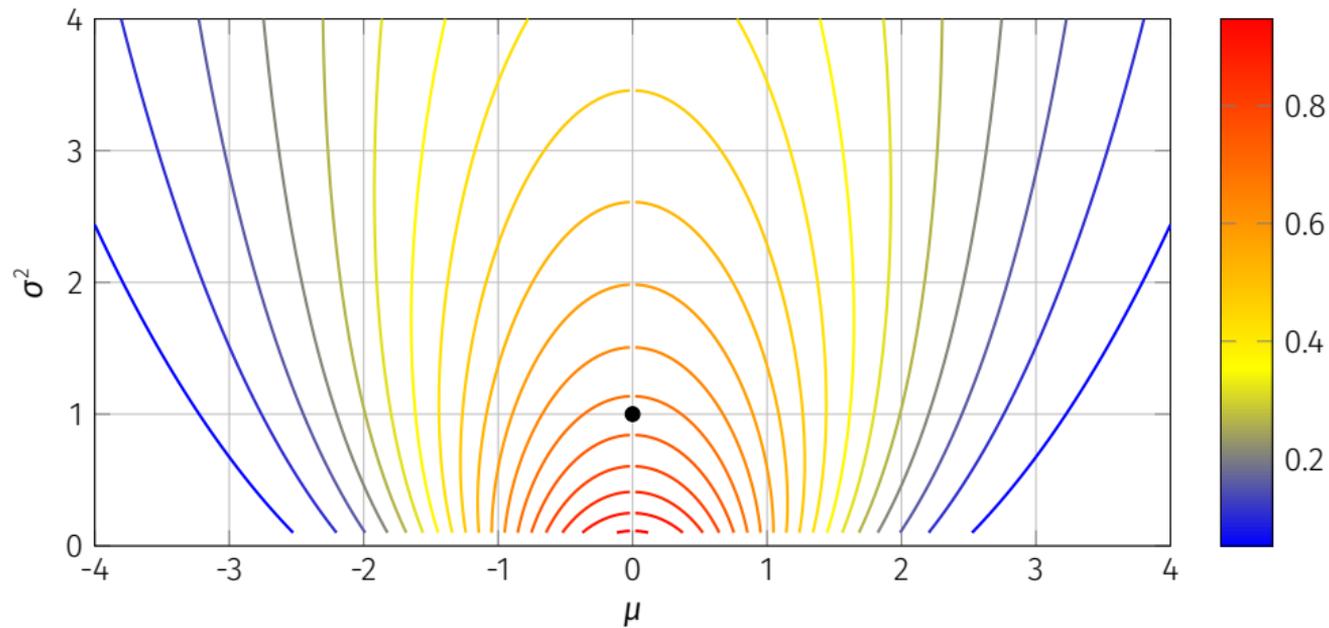
- Une distance entre deux paysages

$$\begin{aligned}
 \|\ell - \ell'\|^2 &= \langle \ell, \ell \rangle + \langle \ell', \ell' \rangle - 2\langle \ell, \ell' \rangle \\
 &= \|\mathbf{p}_\ell - \mathbf{p}_{\ell'}\|_D^2
 \end{aligned}$$

## DISTRIBUTION DU NDVI



## SIMILARITÉ ENTRE CLUSTERS



## EXPÉRIMENTATIONS

EXPÉRIMENTATIONS

MISSION MUESLI

- HySpex (2m/pixel) acquis en Juin 8/9, 2016.

Domaine	Début	Fin
1	400	1340
2	1550	1800
3	1980	2400

- 8 Indices spectraux ont été extraits

NDVI	EVI2	CIRE	NDWI
PIR	VARIG	VARIRE	NDII

- Variables « terrain »

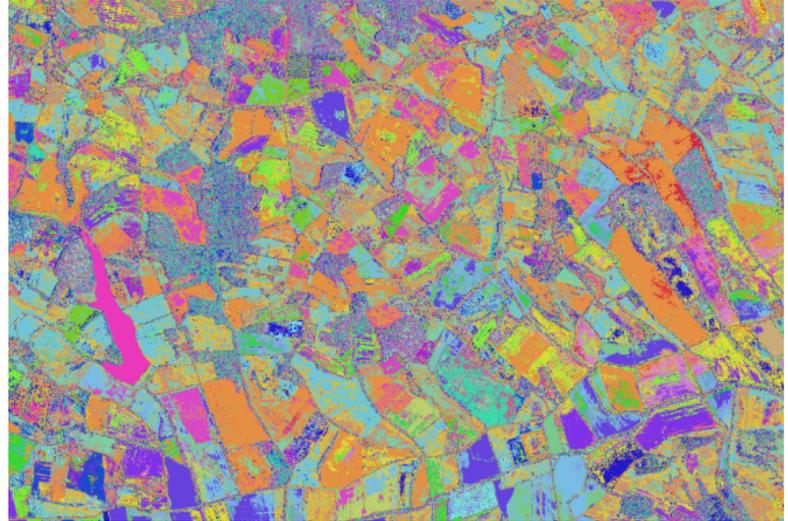
Parcelle agricole	Forêt
Services de production	Services de production
Services de contrôle bio.	Services culturel
Biodiversité floristique	Services de contrôle bio.



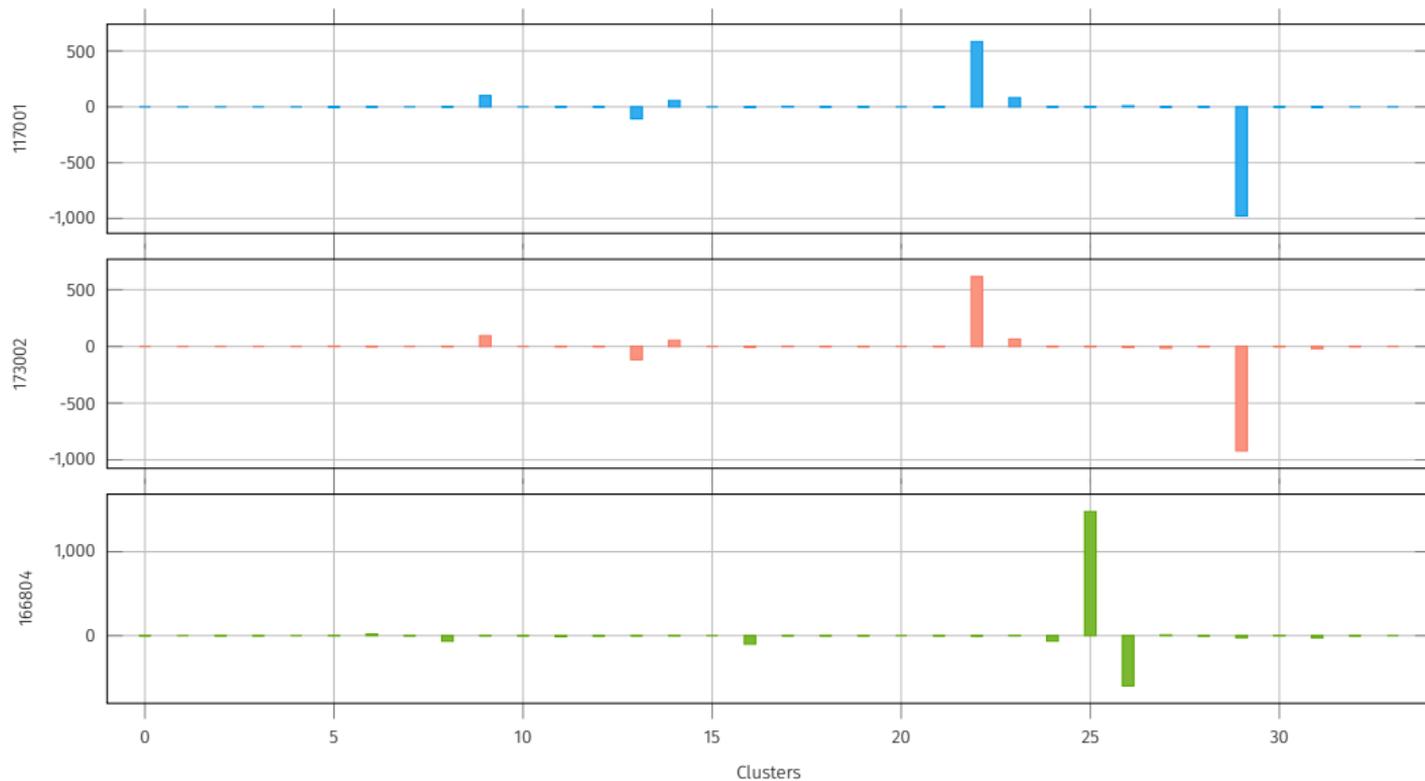
EXPÉRIMENTATIONS

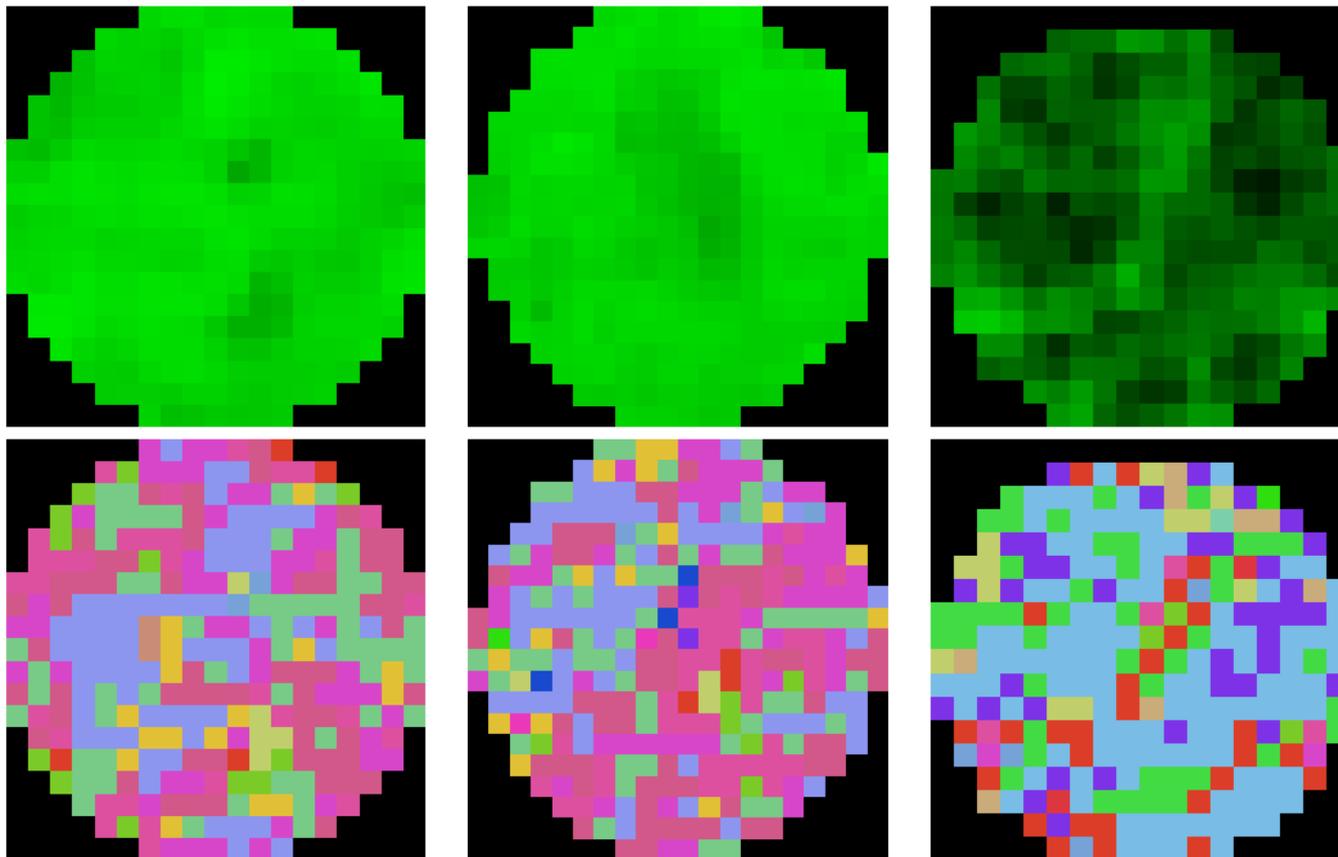
CLUSTERING

## RÉSULTATS "DURS"



## DESCRIPTEURS CONTINUS 1/2





# MATRICE D

	c0	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10	c11	c12	c13	c14	c15	c16	c17	c18	c19	c20	c21	c22	c23	c24	c25	c26	c27	c28	c29	c30	c31	c32	c33
c0	16	8	5	11	0	0	0	3	0	0	10	11	7	10	8	0	3	0	0	0	10	0	0	0	12	0	0	0	8	0	4	0	6	0
c1	8	15	11	1	0	0	0	3	0	0	0	8	11	12	9	0	3	0	0	0	0	0	0	0	7	0	6	0	6	0	9	0	0	0
c2	5	11	13	0	0	0	0	3	0	3	0	5	9	7	8	0	5	0	0	0	0	0	0	6	0	3	0	4	0	6	0	0	0	
c3	11	1	0	11	0	0	0	4	0	0	9	7	0	3	3	0	3	0	0	0	8	0	0	0	8	0	0	5	0	5	0	6	0	
c4	0	0	0	0	7	5	4	0	2	0	0	0	0	0	0	7	0	5	0	0	3	0	0	0	0	1	0	0	0	8	0	0	0	
c5	0	0	0	0	5	13	4	0	0	0	0	0	0	0	0	0	0	13	0	0	0	0	0	0	8	0	0	0	8	0	0	0	0	
c6	0	0	0	0	4	4	11	0	0	0	0	0	0	0	0	6	0	6	0	0	5	0	0	0	0	0	0	0	7	0	0	10	0	
c7	3	3	3	4	0	0	0	4	0	3	4	4	2	3	4	0	3	0	5	4	3	5	3	5	3	0	1	3	5	0	4	0	0	3
c8	0	0	0	0	2	0	0	0	10	0	0	0	0	0	0	6	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	4	6	0
c9	0	0	3	0	0	0	0	3	0	16	0	0	9	0	8	0	12	0	10	8	5	3	2	2	4	0	10	1	11	0	0	0	0	3
c10	10	0	0	9	0	0	0	4	0	0	13	11	0	4	6	0	1	0	0	0	2	0	0	0	7	0	0	0	7	0	6	0	0	0
c11	11	8	5	7	0	0	0	4	0	0	11	15	7	12	10	0	2	0	0	0	0	0	0	8	0	0	0	7	0	9	0	0	0	
c12	7	11	9	0	0	0	0	2	0	9	0	7	12	12	10	0	10	0	5	0	0	0	0	9	0	10	0	9	0	5	0	0	2	
c13	10	12	7	3	0	0	0	3	0	0	4	12	12	16	11	0	5	0	0	0	0	0	0	10	0	4	0	8	0	9	0	0	0	
c14	8	9	8	3	0	0	0	4	0	8	6	10	10	11	11	0	10	0	5	1	0	0	0	10	0	7	0	9	0	6	0	0	3	
c15	0	0	0	0	7	0	6	0	6	0	0	0	0	0	0	12	0	7	0	0	4	0	0	0	0	0	0	0	6	0	0	2	0	0
c16	3	3	5	3	0	0	0	3	0	12	1	2	10	5	10	0	17	0	8	5	6	0	0	0	12	0	9	0	11	0	0	0	0	0
c17	0	0	0	0	5	13	6	0	0	0	0	0	0	0	0	7	0	20	0	0	0	0	0	0	12	0	0	0	12	0	0	0	0	0
c18	0	0	0	0	0	0	0	5	0	10	0	0	5	0	5	0	8	0	15	12	10	12	12	12	3	0	7	11	10	0	0	1	0	11
c19	0	0	0	0	0	0	0	4	0	8	0	0	0	0	1	0	5	0	12	13	9	12	9	8	2	0	4	6	6	0	0	0	0	9
c20	10	0	0	8	3	0	5	3	6	5	2	0	0	0	0	4	6	0	10	9	12	10	6	9	9	0	0	6	8	0	0	7	10	4
c21	0	0	0	0	0	0	0	5	0	3	0	0	0	0	0	0	0	0	12	12	10	17	14	12	0	0	0	11	0	0	0	0	0	12
c22	0	0	0	0	0	0	0	3	0	2	0	0	0	0	0	0	0	0	12	9	6	14	16	13	0	0	0	12	0	0	0	0	0	11
c23	0	0	0	0	0	0	0	5	0	2	0	0	0	0	0	0	0	0	12	8	9	12	13	16	0	0	0	13	0	0	0	1	0	9
c24	12	7	6	8	0	0	0	3	0	4	7	8	9	10	10	0	12	0	3	2	9	0	0	0	16	0	5	0	10	0	1	0	0	0
c25	0	0	0	0	1	8	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0
c26	0	6	3	0	0	0	0	1	0	10	0	0	10	4	7	0	9	0	7	4	0	0	0	5	0	13	0	9	0	0	0	0	0	4
c27	0	0	0	0	0	0	0	3	0	1	0	0	0	0	0	0	0	0	11	6	6	11	12	13	0	0	0	14	0	0	0	0	0	4
c28	8	6	4	5	0	0	0	5	0	11	7	7	9	8	9	0	11	0	10	6	8	0	0	0	10	0	9	0	11	0	2	0	0	1
c29	0	0	0	0	8	8	7	0	0	0	0	0	0	0	0	6	0	12	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0
c30	4	9	6	5	0	0	0	4	0	0	6	9	5	9	6	0	0	0	0	0	0	0	0	1	0	0	0	2	0	13	0	0	0	0
c31	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	1	0	7	0	0	1	0	0	0	0	0	0	0	8	2	0
c32	6	0	0	6	0	0	10	0	6	0	0	0	0	0	0	2	0	0	0	0	10	0	0	0	0	0	0	0	0	0	2	15	0	0
c33	0	0	0	0	0	0	0	3	0	3	0	0	2	0	3	0	0	0	11	9	4	12	11	9	0	0	4	4	1	0	0	0	0	15

EXPÉRIMENTATIONS  
ANALYSE PRÉDICTIVE

- 4 représentations paysagères
  1. Distance moyenne au centroid
  2. Moments
  3. Variance inter/intra
  4. Continue
- 4 algorithmes de prédiction
  1. Régression Ridge
  2. K-NN
  3. RF
  4. Kernel-RR avec sélection de variables
- LOOCV, avec standardisation des x et y
- Qualité de la prédiction :  $r^2$  moyen

- 4 représentations paysagères

1. Distance moyenne au centroid
2. Moments
3. Variance inter/intra
4. Continue

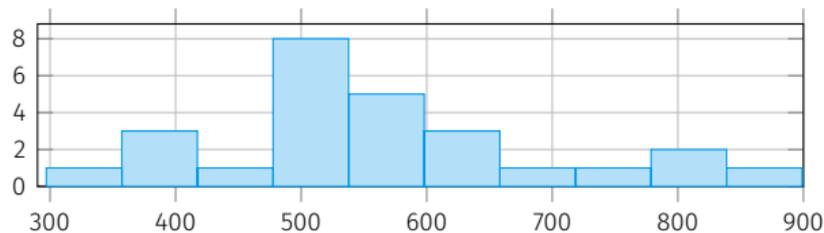
- 4 algorithmes de prédiction

1. Régression Ridge
2. K-NN
3. RF
4. Kernel-RR avec sélection de variables

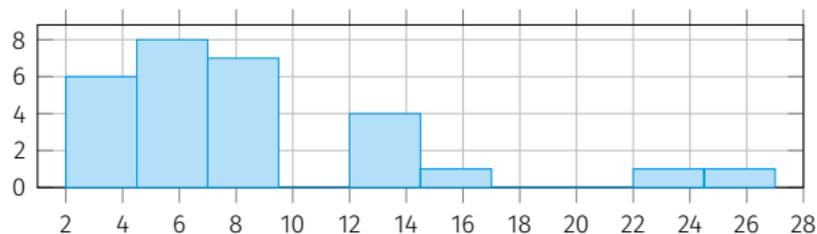
- LOOCV, avec standardisation des x et y

- Qualité de la prédiction :  $r^2$  moyen

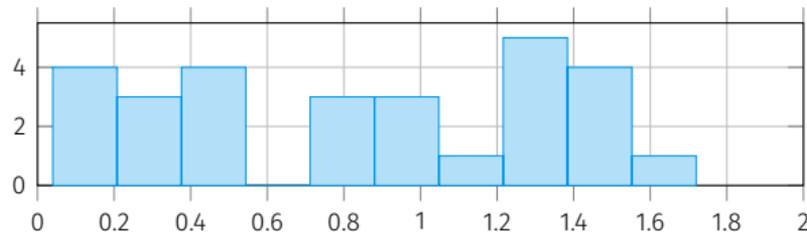
Nombre d'épis



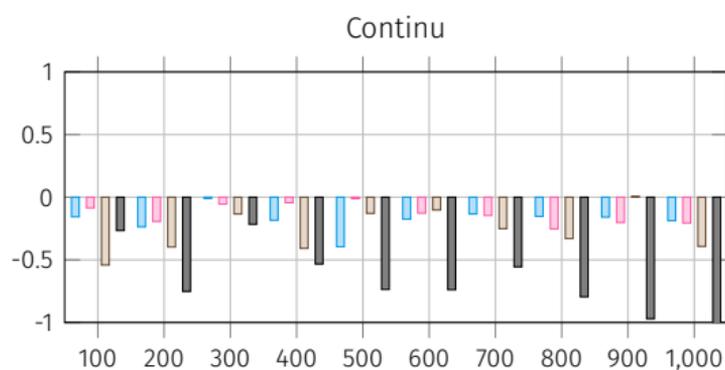
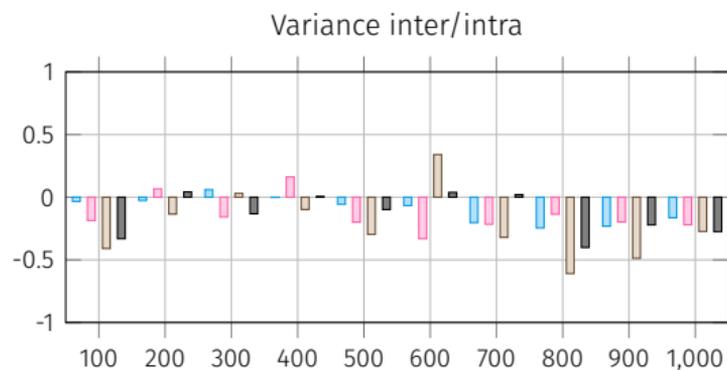
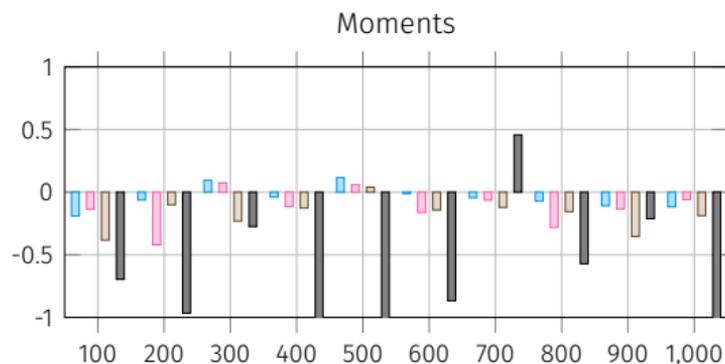
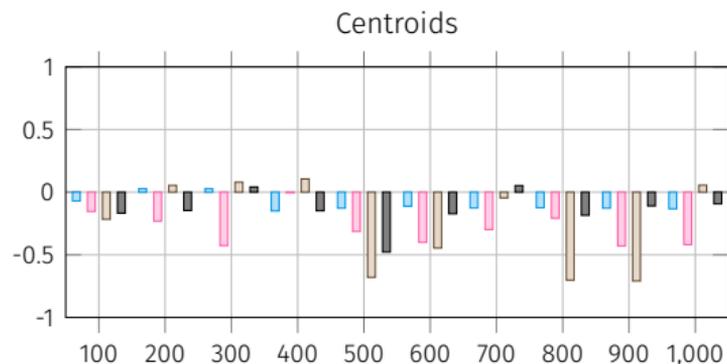
Nombre d'espèces



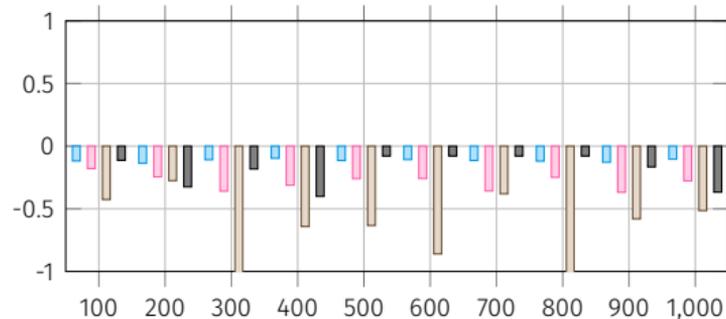
Indice de Shannon



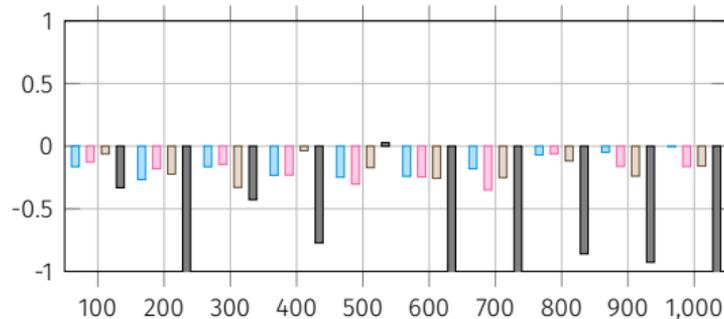
# NOMBRE D'ÉPIS



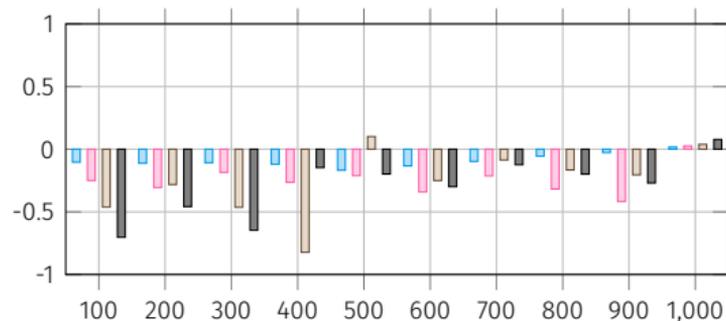
## Centroids



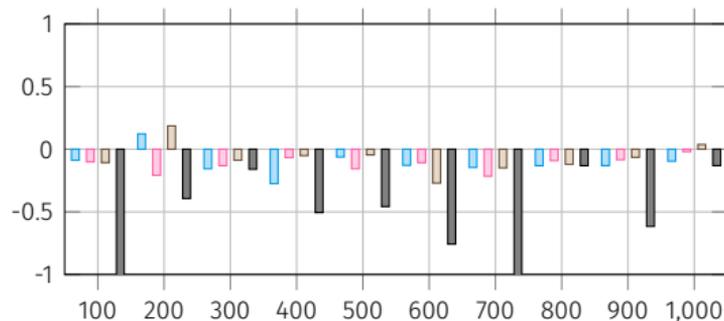
## Moments



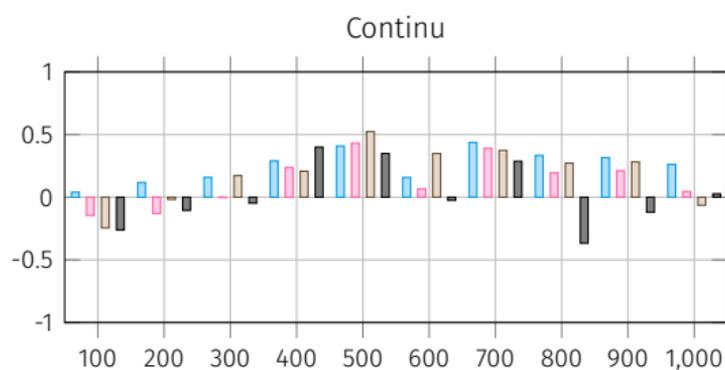
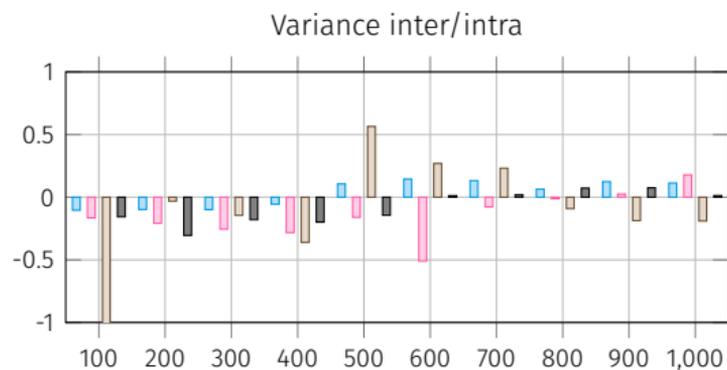
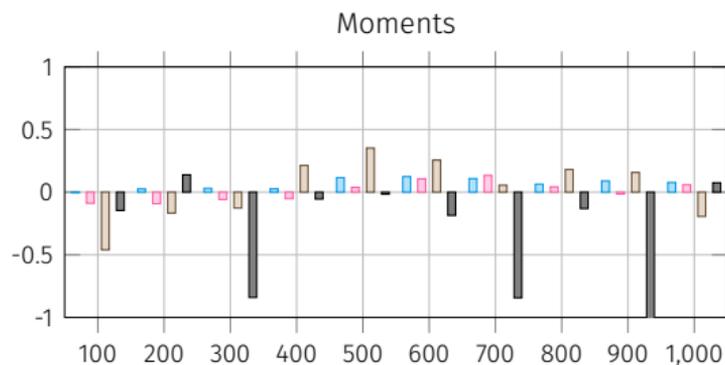
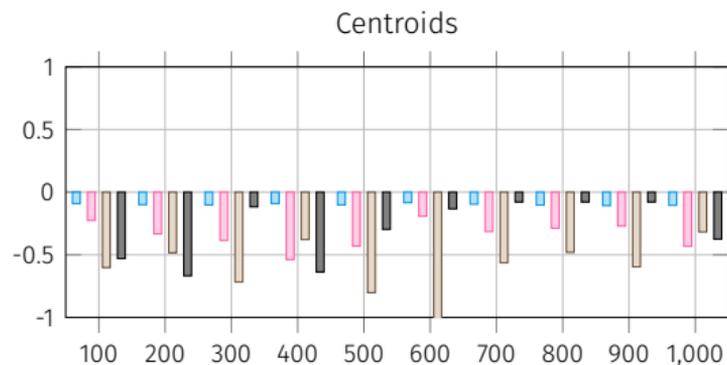
## Variance inter/intra

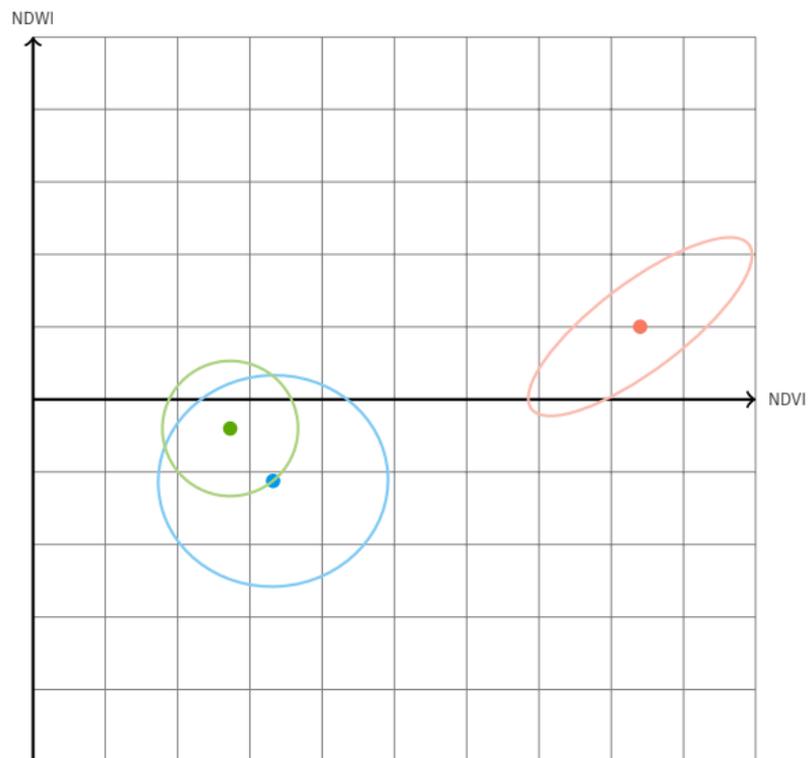


## Continu



# NOMBRE D'ESPÈCES





## CONCLUSION ET PERSPECTIVES

### Approche continue

- Modélisation par GMM de la radiométrie
- Modélisation des paysages par la distribution locale
- Définition d'un produit scalaire

### Résultats

- Meilleurs résultats pour l'approche continue sur des mesures de richesses
- **Mais**
  - ▶ Faible jeu de données
  - ▶ Que des données de réflectance

- Influence du nombre de classes ?
- D'autres descripteurs que la distribution conditionnelle ?
- Inclusion d'autres données (LiDar, Radar, multi-temporel) ?

DES QUESTIONS/REMARQUES ?