

# Optimisation de configuration spectrale pour la classification – exemple d'application au milieu urbain

A. Le Bris (IGN),  
N. Chehata (UMR LISAH – IRD & ENSEGID)  
X. Briottet (ONERA)  
N. Paparoditis (IGN)

# Optimisation de configuration spectrale pour la classification – exemple d'application au milieu urbain

# Cartes de matériaux

**Des besoins pour une occupation du sol très fine en milieu urbain,**  
→ **très grande échelle + sémantique**  
**i.e. pour des cartes de matériaux**

- Quantification des flots de polluants émanant des matériaux de toiture et emportés par ruissellement
- Détection de matériaux dangereux (ex : fibro-ciment)
- Surveillance de l'état des routes
- Distinction entre surfaces perméables et imperméables
  - Modèles hydrologiques
  - Suivi de l'extension des surfaces artificialisées / conformité à la réglementation
- Donnée de base pour des simulateurs :
  - modèle climatique très local
  - propagation d'ondes électro-magnétique

**Quelles sont les bandes spectrales les plus utiles pour ce type d'application ?**

→ sélection d'un ensemble optimal de bandes spectrales (position+largeur)  
vers une caméra superspectrale dédiée à ces applications.

# Objectif visé : une caméra superspectrale

## Quelles sont les bandes spectrales les plus utiles pour ce type d'applications ?

- sélection d'un ensemble optimal de bandes spectrales (position+largeur)  
vers une caméra superspectrale dédiée à ces applications.

## Caméra IGN modulaire

### Avantages potentiels d'une telle caméra :

- meilleure résolution spatiale
- uniquement bandes spectrales utiles pour le problème donné
- fauchée plus large
- caméra matricielle
  - possibilité d'acquisition multivue
  - visées multiples / BRDF
- bandes spectrales plus larges → plus de photons

# Objectif visé : une caméra superspectrale

**Comment identifier les bandes les plus pertinentes ?**

**Ou du moins identifier les parties du spectre les plus utiles ?**

→ **Mise en œuvre d'une méthodologie de sélection d'attributs**

**2 possibilités :**

- **Travail à partir d'une base de données de spectres de référence**
- **Travail à partir d'acquisitions hyperspectrales aéroportées THR sur des zones représentatives**
  - Permet d'avoir des matériaux présentant des aspects variable
  - Conditions plus proches de celles la future caméra
  - Possibilité de simuler différentes résolutions spatiales

# Objectif visé : une caméra superspectrale

**Comment identifier les bandes les plus pertinentes ?**

**Ou du moins identifier les parties du spectre les plus utiles ?**

→ **Mise en œuvre d'une méthodologie de sélection d'attributs**

**2 possibilités :**

- **Travail à partir d'une base de données de spectres de référence**
- **Travail à partir d'acquisitions hyperspectrales aéroportées THR sur des zones représentatives**
  - Permet d'avoir des matériaux présentant des aspects variable
  - Conditions plus proches de celles la future caméra
  - Possibilité de simuler différentes résolutions spatiales

## **Campagne d'acquisition UMBRA (IGN + ONERA) sur Toulouse et Amiens**

**Des paysages urbains différents :**

- différents types de quartiers : centre-villes denses, zones résidentielles d'habitat individuel ou collectif, zones industrielles et commerciales, infrastructure de transport (autoroute, ferroviaire, aéroport), zones végétalisés (parcs, jardins) et milieu aquatique (rivières, étangs)...

- 2 villes pouvant être considérées comme typiques de bon nombre de villes françaises :  
Amiens pour les villes du Nord et Toulouse pour le Sud-Ouest.  
→ grande variété de matériaux urbains



# La campagne d'acquisition UMBRA

**Acquisition d'images aériennes sur Toulouse et Amiens**

**Paysages urbains différents → plus grande diversité de matériaux**

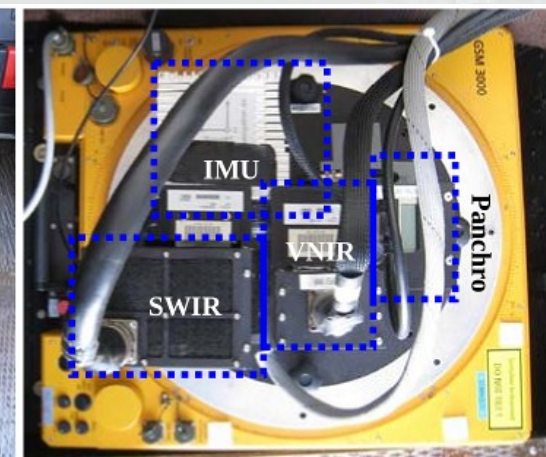
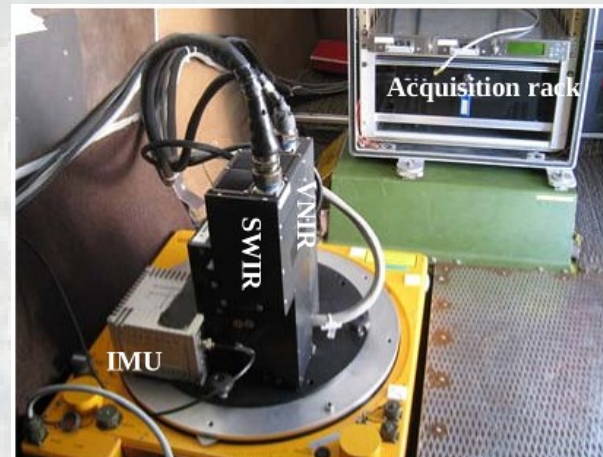
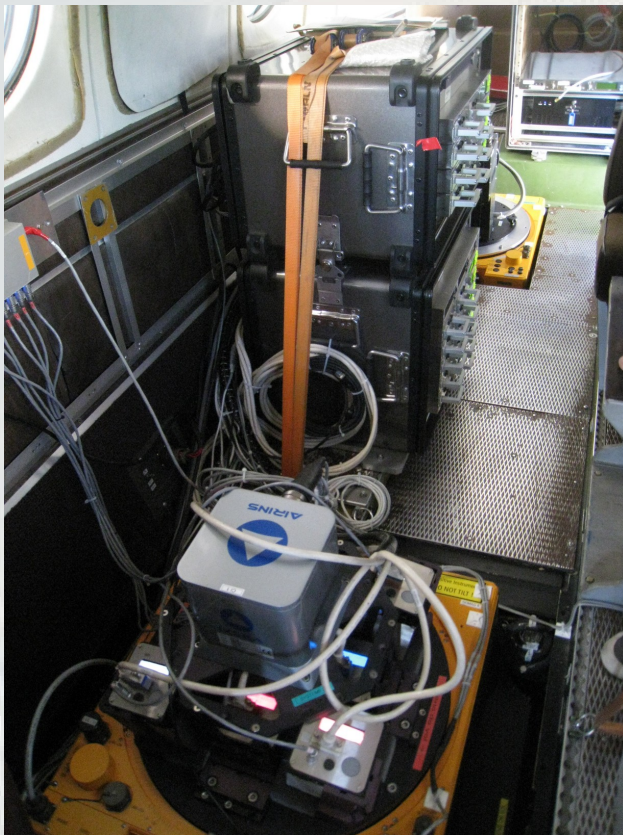
**Des capteurs différents à bord du même avion :**

- très haute résolution spectrale → **hyperspectral Hypspx (VNIR+SWIR)**
- très haute résolution spatiale → **CamV2 IGN**
- données de télédétection et photogrammétriques

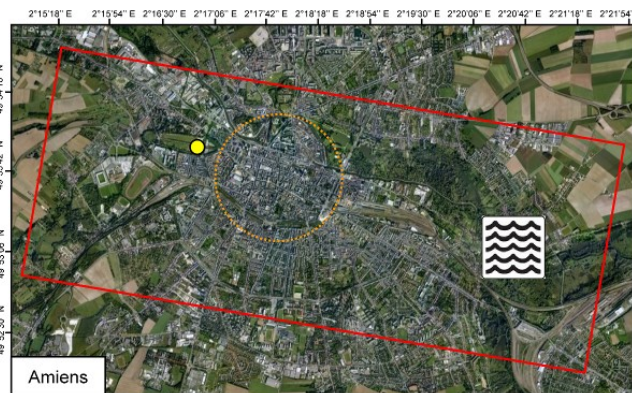
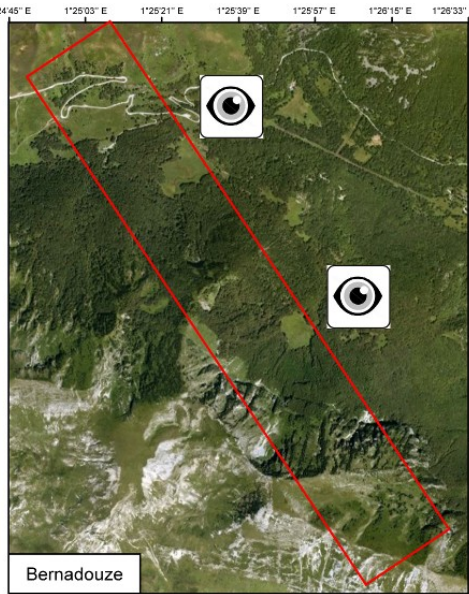
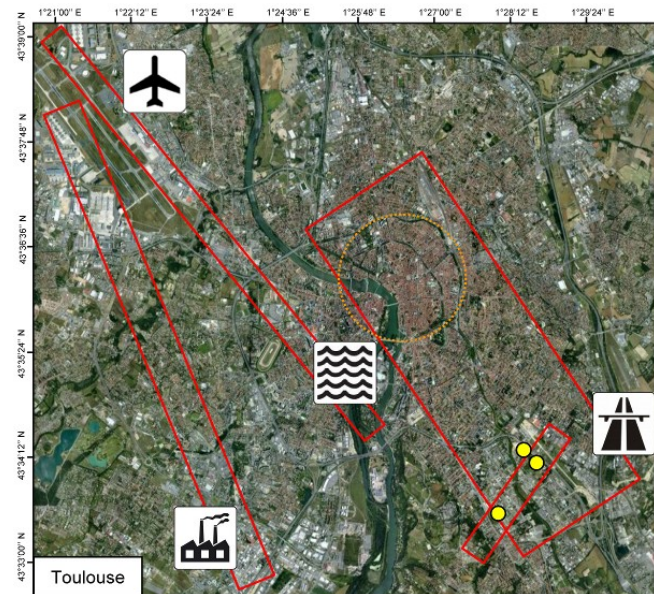
[Adeline,2013]

**Différentes mesures radiométriques au sol**







Description	Système PELICAN-HYPER			Système IGN	
	HySpex VNIR	HySpex SWIR	CamV2 panchro	CamV2 multi	CamV2 panchro
Type d'acquisition	barrette	barrette	matrice	matrice	matrice
Nombre de pixels	1600	320	7256 x 5462	7256 x 5462	7256 x 5462
Intervalle spectral	0,4-1 $\mu\text{m}$	1-2,5 $\mu\text{m}$	0,4-0,8 $\mu\text{m}$	0,4-0,8 $\mu\text{m}$	0,35-0,75 $\mu\text{m}$
Nombre de bandes spectrales	160	256	panchro	4	panchro
Résolution spatiale	0,80m	1,6m	0,14m	0,24m	0,12m
Fauchée	645m	520m	1033m	1775m	1775m

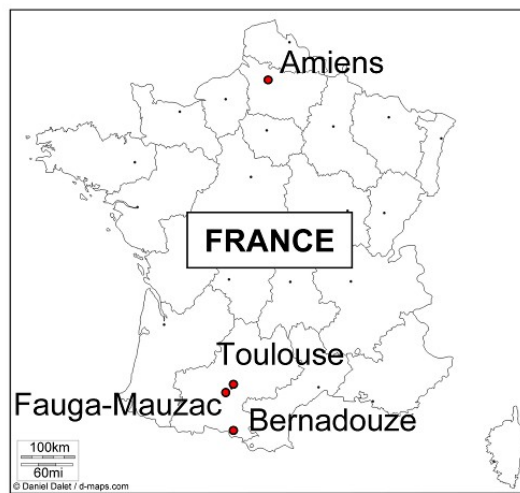






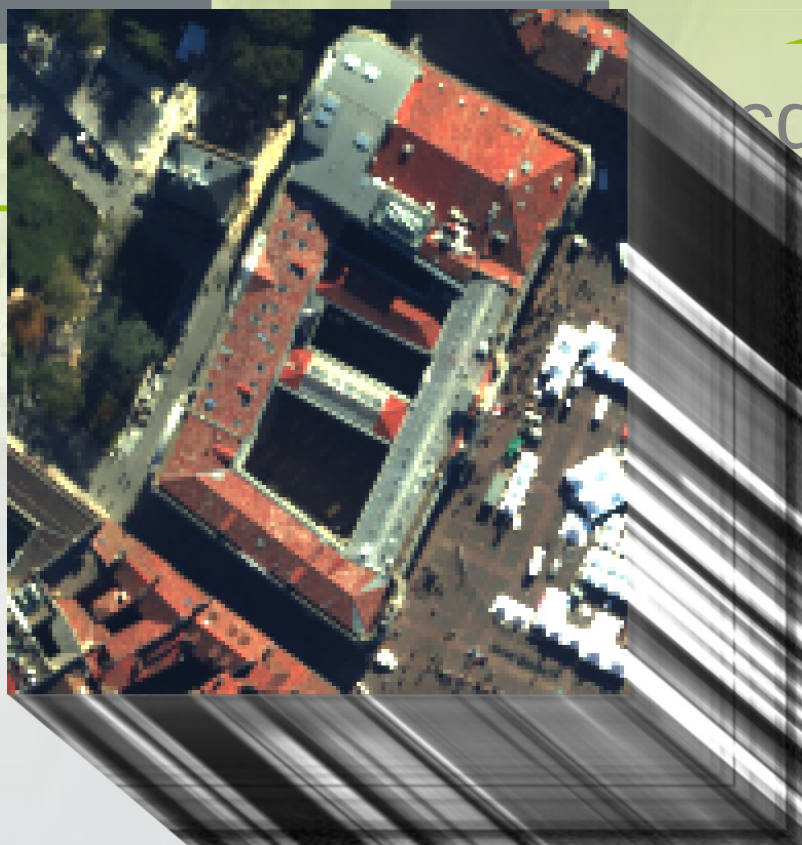
#### Légende:

-  Vieillessement des pistes de vol aéroportuaires
-  Qualité de l'air et pollution industrielle
-  Pollution urbaine (autoroutes)
-  Problématiques hydrologiques
-  Contrôle des zones brûlées
-  Observation de l'impact de l'homme sur l'environnement montagneux



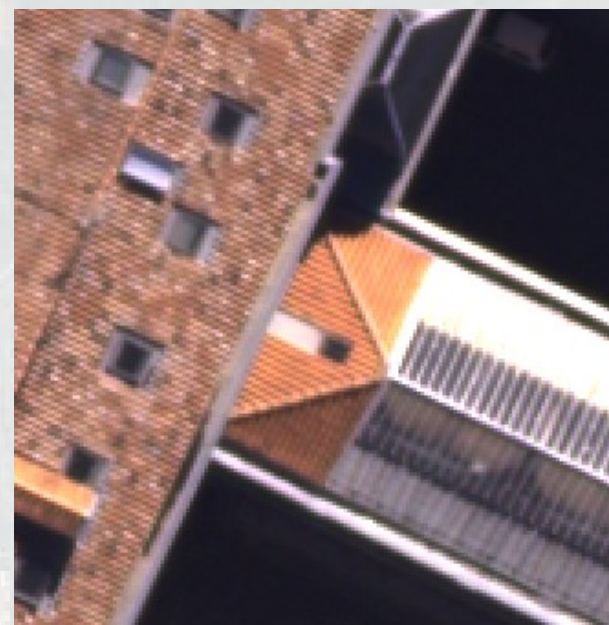
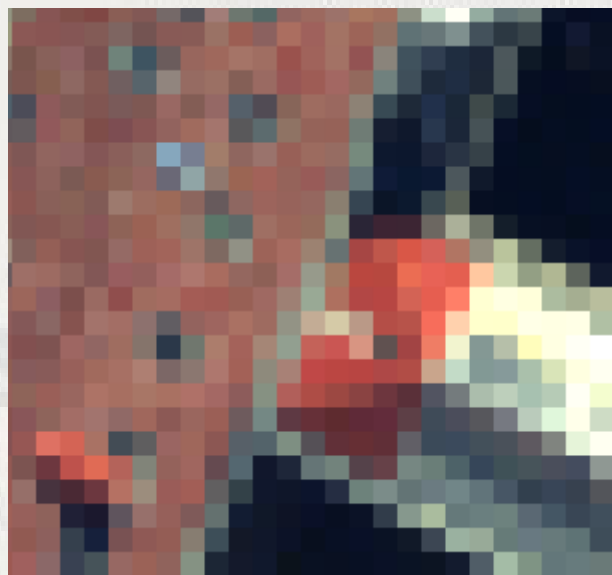


acquisition



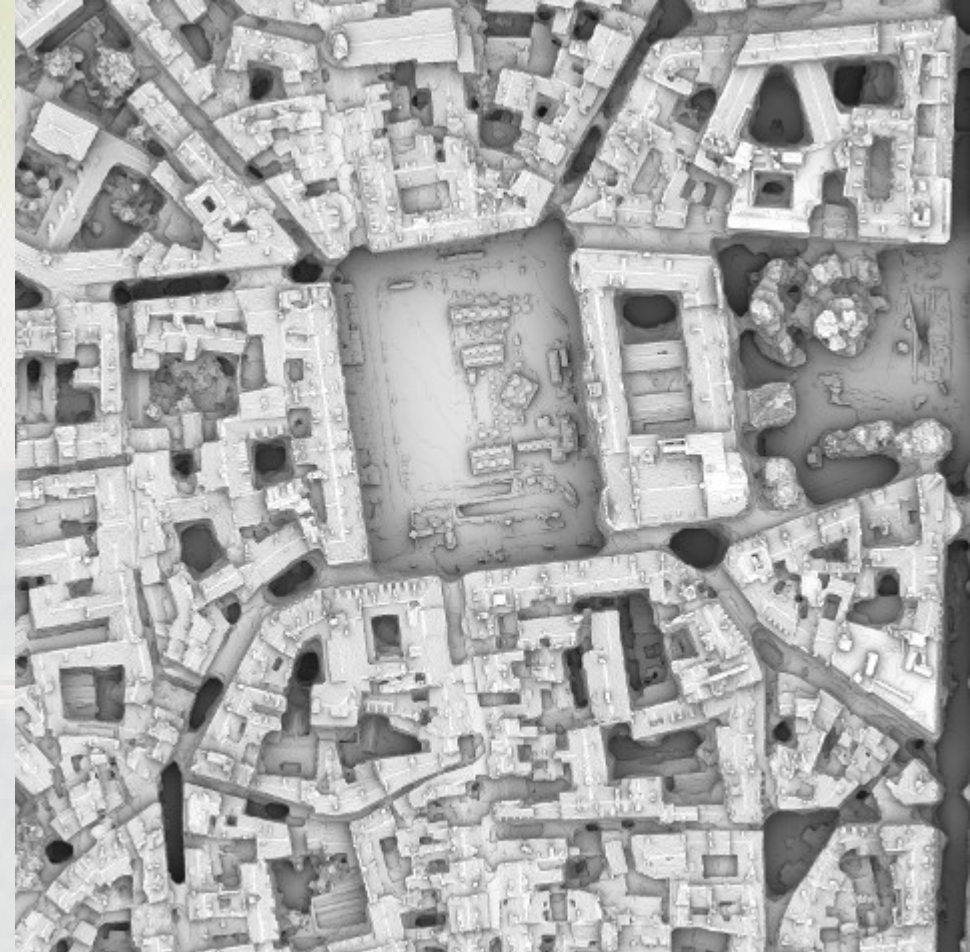
HySpex VNIR : 160 bandes / pixel sol 80cm  
HySPex SWIR : 256 bandes / pixel sol 1,6 m

CamV2 : 4 bandes / pixel sol 12 cm





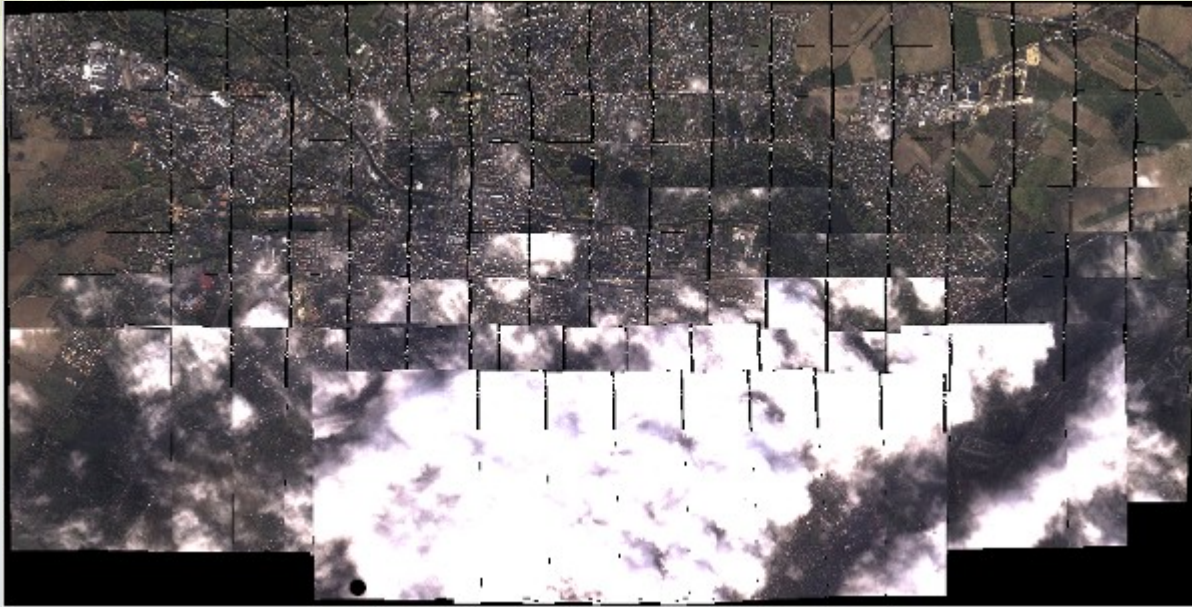
# La campagne d'acquisition UMBRA



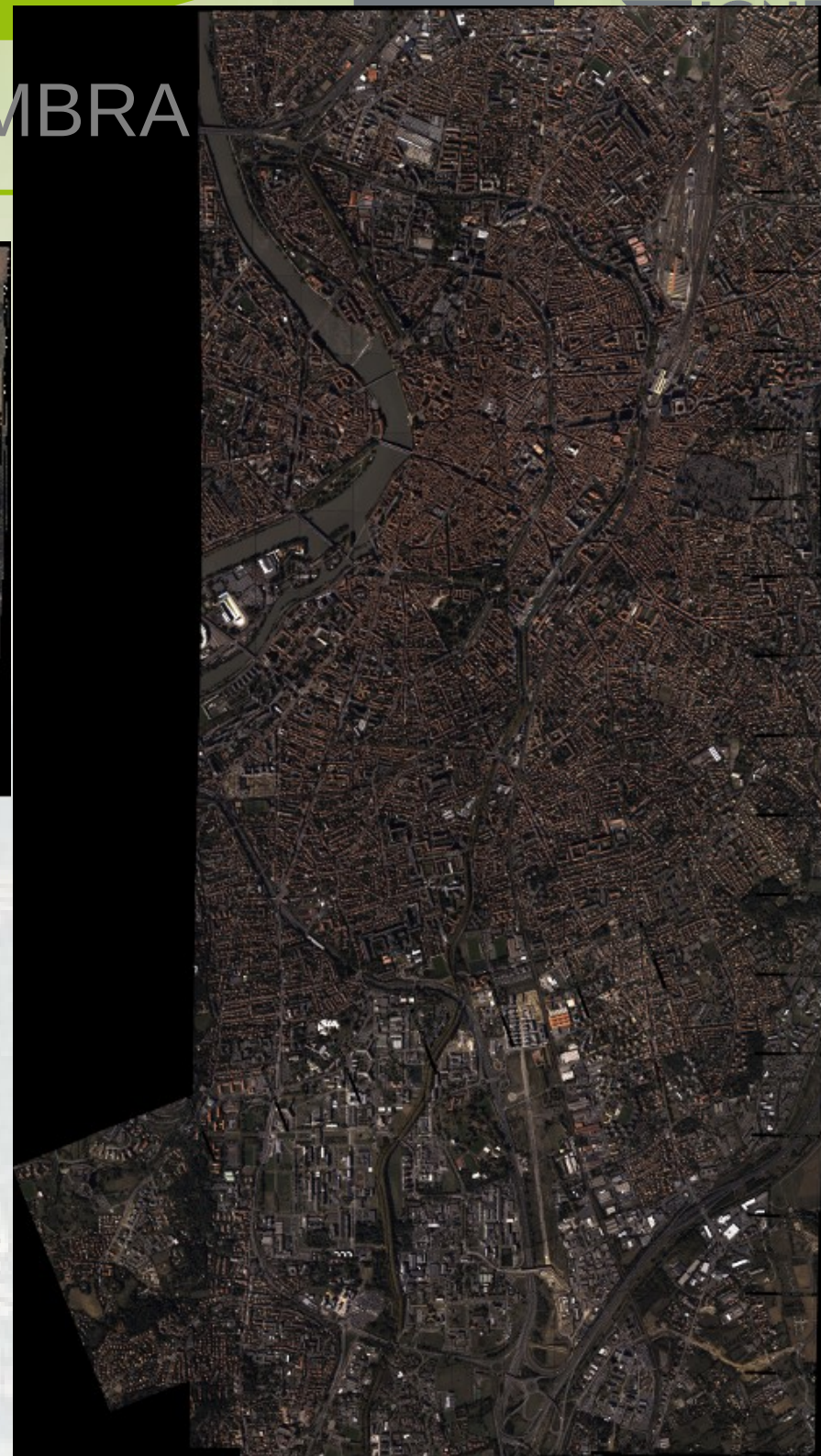
MNS de corrélation



# La campagne d'acquisition UMBRA



Amiens



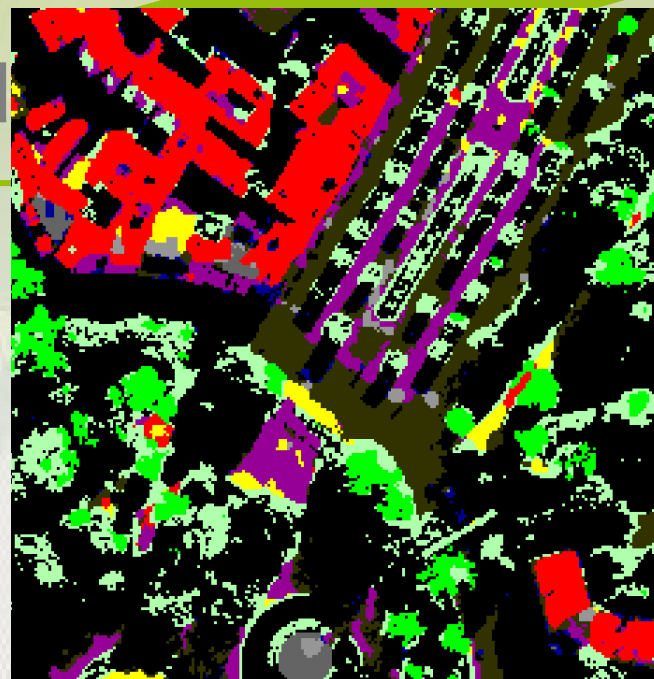
Toulouse



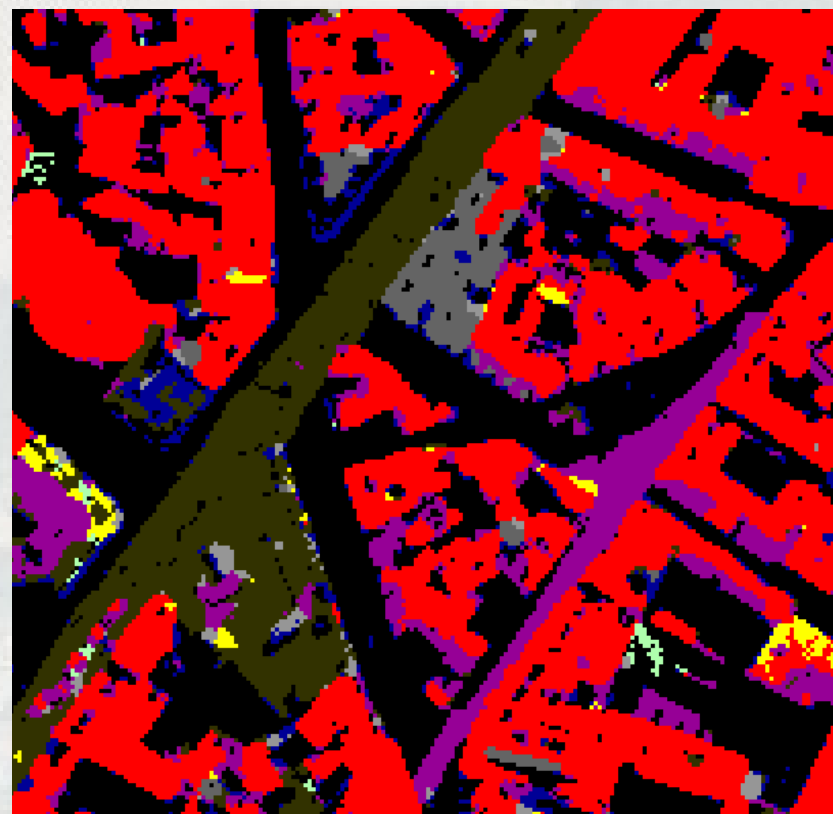
# Premiers tests de classification



- Tuiles rouges
- Métal 1
- Métal 2
- Ardoises
- Pavés
- Asphalté
- Sol nu
- Végétation haute
- Végétation basse



- Tuiles rouges
- Métal 1
- Métal 2
- Ardoises
- Pavés
- Asphalte
- Sol nu
- Végétation haute
- Végétation basse



# Sélection de bandes

**Un problème déjà bien abordé dans la littérature...**

**Sélection de bandes = critère/score à optimiser + méthode d'optimisation**

- **“Filter”**

RELIEFF [Kononenko,1997]

Importances calculées à parti d'ACP [Chang,1999] ou d'ICA [Du,2003]

Mesures de séparabilité : distance de Bhattacharyya ou Jeffries-Matusita [Bruzzone,2000][Herold,2004]  
[Serpico,2007], ou MEAC [Yang,2012]

Information mutuelle [Battiti,1994][Martinez-Usó,2007][Guo,2008][Estevez,2009][Le Moan,2011][Cang,2012]

“Orthogonal Projection divergence” [Du,2008]

- **“Wrapper”** → met en oeuvre un classifieur

SVM [Estevez,2009][Li,2010][Yang,2007]

Random Forests [Diaz-Uriarte,2006]

ML [Zhang,2007]

- **“Embedded”**

SVM-RFE [Guyon,2002], MKL-SVM [Tuia,2010]

Importance issue des Random Forests [Breiman,2001]

Marges des SVM [Fröhlich,2003] ou des RF [Pal,2009]

- **Approches hybrides**



# Sélection de bandes

**Un problème déjà bien abordé dans la littérature...**

**Sélection de bandes = critère/score à optimiser + méthode d'optimisation**

- "Filter"
- "Wrapper"
- "Embedded"

## Méthodes d'optimisation

- Méthodes séquentielles

Sequential Forward Search (SFS), Sequential Backward Search (SBS) et leurs variantes flottantes  
SFFS et SBFS [Pudil,1994]

Steepest Ascent [Serpico,2001]

- Méthodes stochastiques

Algorithmes génétiques [Estevez,2001][Li,2011][Zhuo,2008][Minet,2010][Fröhlich,2003]

Particle Swarm Optimization (PSO) [De Backer,2005][Yang,2007]

Clonal selection [Zhang,2007]

Colonie de fourmis [Zhou,2009]

Recuit simulé [Chang,2011]

## Sequential Forward Floating Search (SFFS) → séquentiel

- Initialisation : recherche de la bande  $b$  qui permet d'obtenir le meilleur score  

$$S \leftarrow \{b\}$$

Ajout d'une nouvelle bande à la sélection jusqu'à critère d'arrêt

- Recherche de la bande  $b$  qui associée aux bandes sélectionnées permet d'obtenir le meilleur score.

$$S \leftarrow S \cup \{b\}$$

- Remise en cause des bandes sélectionnées : recherche de la bande  $s$  dont la suppression dégrade le moins le score

si  $s=b$ , on passe au niveau suivant.

$$\text{Sinon } S \leftarrow S \setminus \{s\}$$

A chaque niveau, on dispose d'un ensemble bandes sélectionnées proche de l'optimum...

→ Permet de connaître l'impact du nombre de bandes sur le résultat.

# Algorithme génétique

## Algorithme génétique (AG) → stochastique

- utilisé ici pour rechercher un nombre  $n$  fixe de bandes fourni en paramètre
- définitions :  
un individu = un ensemble de  $n$  bandes  $\{b_1 \dots b_n\}$   
une population = un ensemble d'individus
- Initialisation : Génération aléatoire d'une population initiale  $P_0$  de  $N_{ini}$  individus

Itération  $i$  :

- Calcul du score associé à chaque individu de la population  $P_i$
- Tri des individus de la population  $P_i$  en fonction de ce score
- On ne conserve que les  $N$  meilleurs individus de  $P_i$ . Soit  $S_i$  cette population de survivants.

- Création d'une nouvelle génération

Pour chaque nouvel individu :

sélection aléatoire de 2 parents dans  $S_i$

le nouvel individu est issu du croisement aléatoire de ces parents

des mutations aléatoires se produisent également pour éviter les optima locaux



# Algorithme génétique

## Algorithme génétique (AG) → stochastique

- utilisé ici pour rechercher un nombre  $n$  fixe de bandes fourni en paramètre
- définitions :  
un individu = un ensemble de  $n$  bandes  $\{b_1 \dots b_n\}$   
une population = un ensemble d'individus
- Initialisation : Génération aléatoire d'une population initiale  $P_0$  de  $N_{ini}$  individus

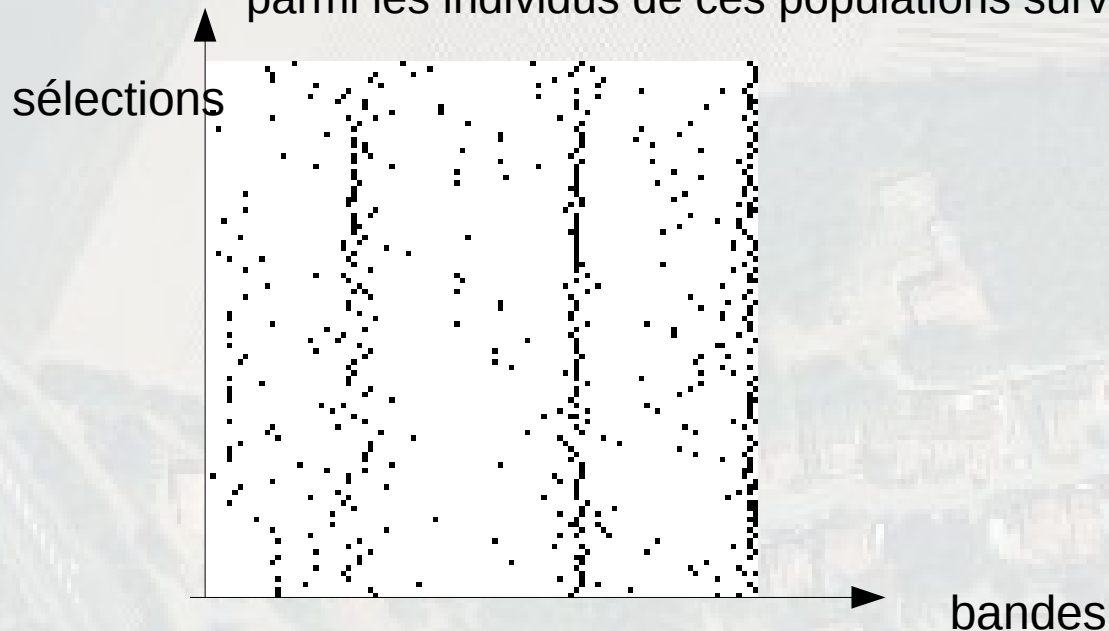
Itération  $i$  :

- Calcul du score associé à chaque individu de la population  $P_i$
  - Tri des individus de la population  $P_i$  en fonction de ce score
  - **On ne conserve que les  $N$  meilleurs individus de  $P_i$ . Soit  $S_i$  cette population de survivants.**
  - Création d'une nouvelle génération
- Pour chaque nouvel individu :
- sélection aléatoire de 2 parents dans  $S_i$
  - le nouvel individu est issu du croisement aléatoire de ces parents
  - des mutations aléatoires se produisent également pour éviter les optima locaux

# Algorithme génétique

## Algorithme génétique (AG) → stochastique

- utilisé ici pour rechercher un nombre  $n$  fixe de bandes fourni en paramètre
- **On ne conserve que les  $N$  meilleurs individus de  $P_i$ . Soit  $S_i$  cette population de survivants.**
- Mais, il n'y a pas une bonne solution et toutes les autres mauvaises...  
→ Les populations survivantes contiennent aussi des bonnes solutions
- Calcul pour chaque bande d'un score correspondant à sa fréquence de présence parmi les individus de ces populations survivantes.



# Algorithme génétique

## Algorithme génétique (AG) → stochastique

- utilisé ici pour rechercher un nombre  $n$  fixe de bandes fourni en paramètre
- **On ne conserve que les  $N$  meilleurs individus de  $P_i$ . Soit  $S_i$  cette population de survivants.**
- Mais, il n'y a pas une bonne solution et toutes les autres mauvaises...
  - Les populations survivantes contiennent aussi des bonnes solutions
- Calcul pour chaque bande d'un score correspondant à sa fréquence de présence parmi les individus de ces populations survivantes.



# Algorithme génétique

## Algorithme génétique (AG) → stochastique

- utilisé ici pour rechercher un nombre  $n$  fixe de bandes fourni en paramètre
- **On ne conserve que les  $N$  meilleurs individus de  $P_i$ . Soit  $S_i$  cette population de survivants.**
- Mais, il n'y a pas une bonne solution et toutes les autres mauvaises...
  - Les populations survivantes contiennent aussi des bonnes solutions
  - Calcul pour chaque bande d'un score correspondant à sa fréquence de présence parmi les individus de ces populations survivantes.

En pratique :

- on lance l'algorithme génétique sur plusieurs sous-jeux de données
- on lance plusieurs fois l'algorithme génétique → cf initialisations différentes
  - on calcule alors la moyenne et l'écart-type des scores obtenus

# Scores à optimiser

- **Scores liés à un classifieur**

2 classifieurs testés

- SVM linéaire un-contre-un sans optimisation des paramètres
- Forêts aléatoires (RF)

- Scores "taux de bonne classification"

- Précision globale
- Précision moyenne
- Kappa
- Fscore moyen
- Fscore minimal

- Scores "marge"

- Marge d'un classifieur SVM
- Marge d'un classifieur Forêts aléatoires (RF)

$$score = \sum_p \delta(p, RF(p)) \cdot probaRF(p)$$

$$\delta(i, j) = 1 \text{ si } i = j$$

$$\delta(i, j) = -1 \text{ sinon}$$

# Scores à optimiser

- **Scores liés à un classifieur**

2 classifieurs testés

- SVM linéaire un-contre-un sans optimisation des paramètres
- Forêts aléatoires (RF)

- Scores "taux de bonne classification"

- Scores "marge"

- **Scores "séparabilité" entre classes**

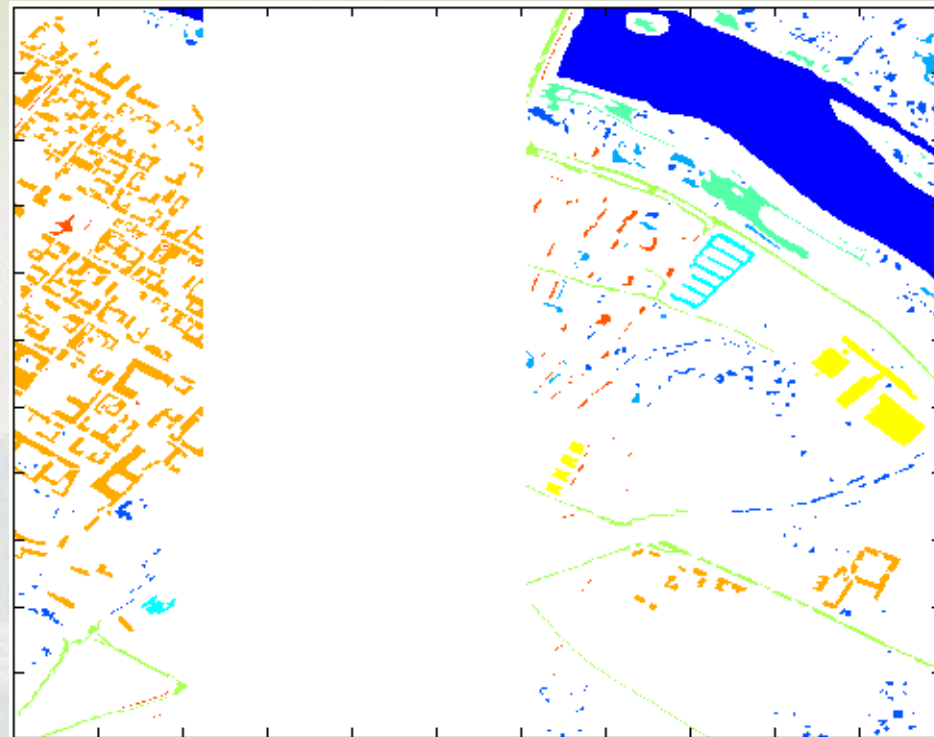
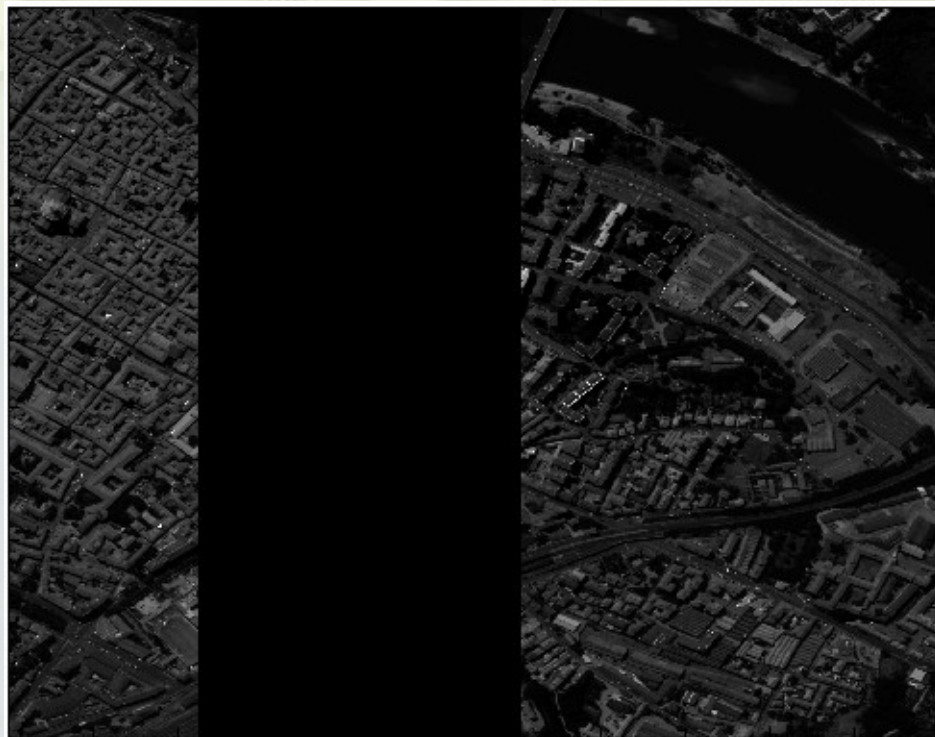
- distance de Bhattacharyya (BDistance) moyenne ou minimale
- mesure de Jeffries-Matsuhita

- **Scores "information mutuelle"**

- compromis entre corrélation avec la vérité terrain et non redondance entre bandes sélectionnées [Estevez, 2009]

$$\sum_{s \in S} I(s, C) - \sum_{(i \in S ; i \neq s)} \frac{I(s, i)}{\min(H(s), H(i))}$$

# Jeu test Pavia



## Légende :

- Eau
- Arbres
- Pelouse
- Pavés
- Terre / sol nu
- Asphalte
- Toiture goudronnée
- Tuiles
- Ombres

Données ROSIS

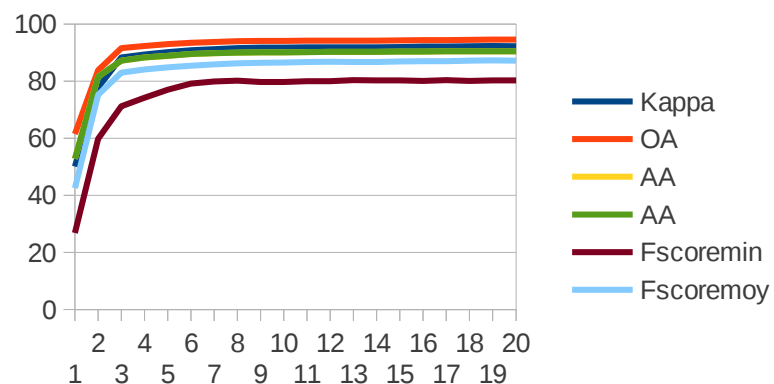


# Jeu test Pavia

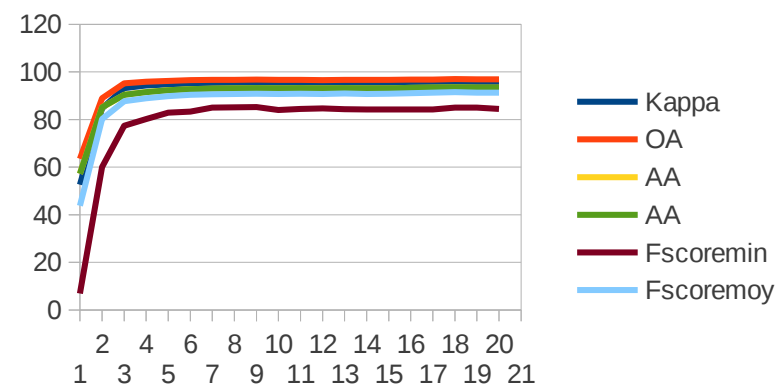
Evolution d'indicateurs de bonne classification en fonction du nombre de bandes sélectionnées par SFFS

score à optimiser = séparabilité Bdist moyenne

Évaluation par classification RF

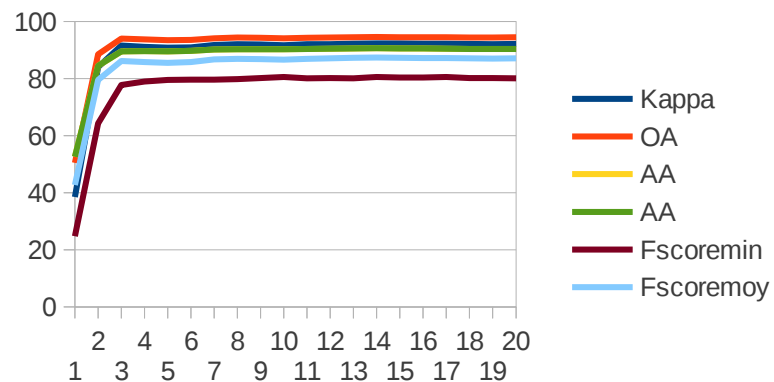


Évaluation par classification SVM gaussienne (avec optimisation)

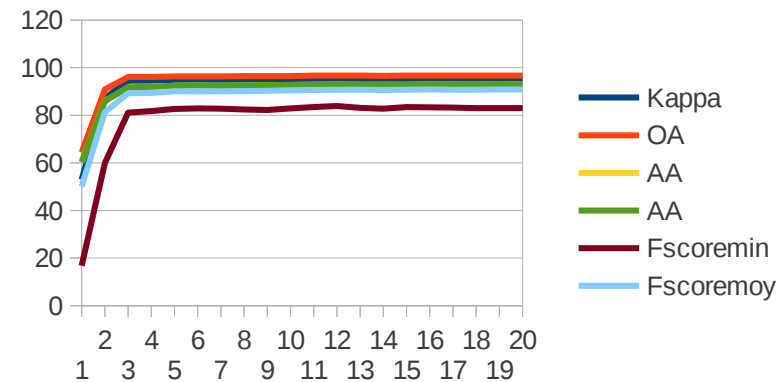


score à optimiser = précision moyenne d'un classification SVM linéaire

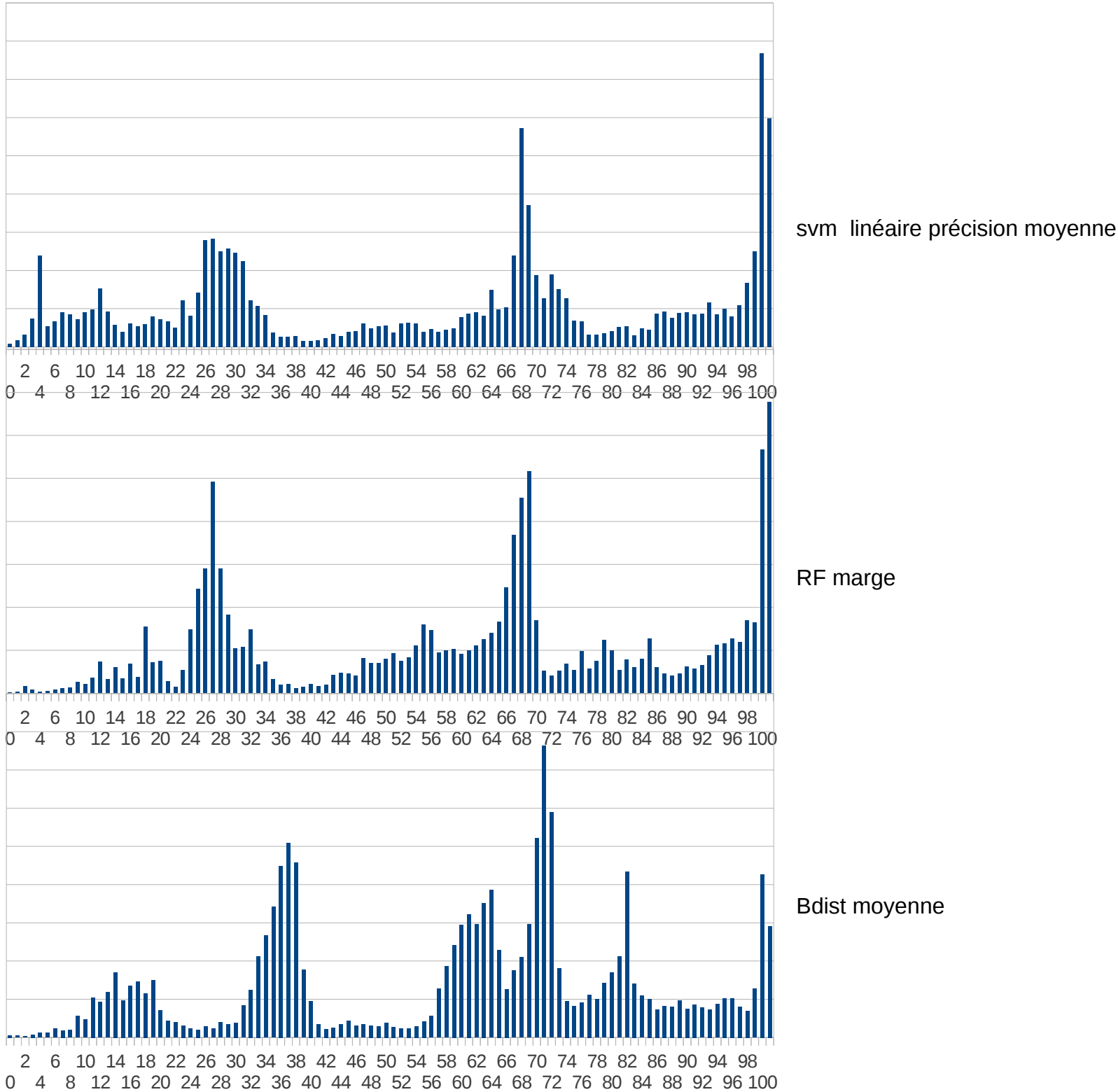
Évaluation par classification RF



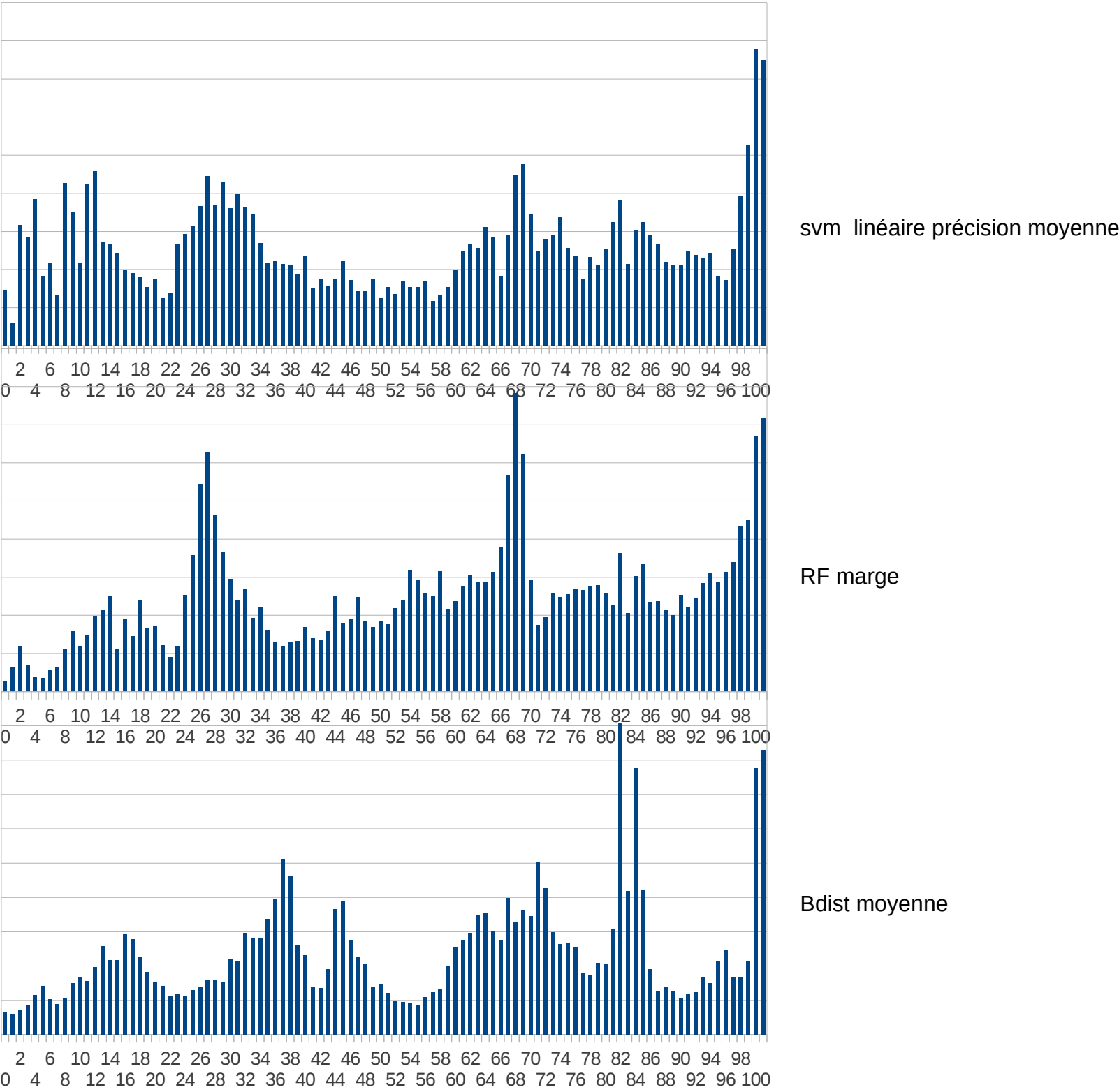
Évaluation par classification SVM gaussienne (avec optimisation)



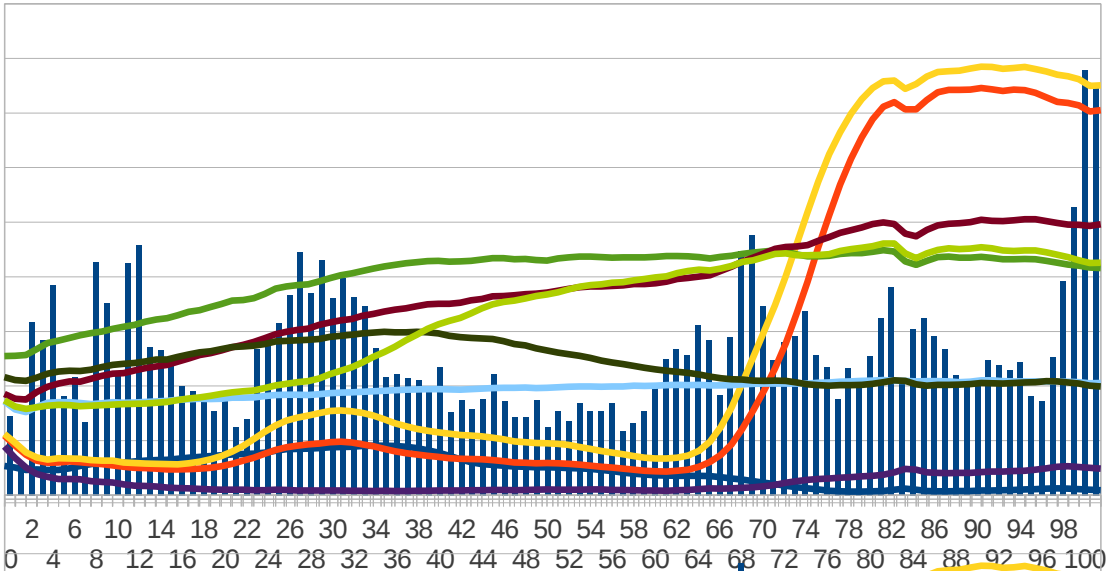
Sélection de 5 bandes



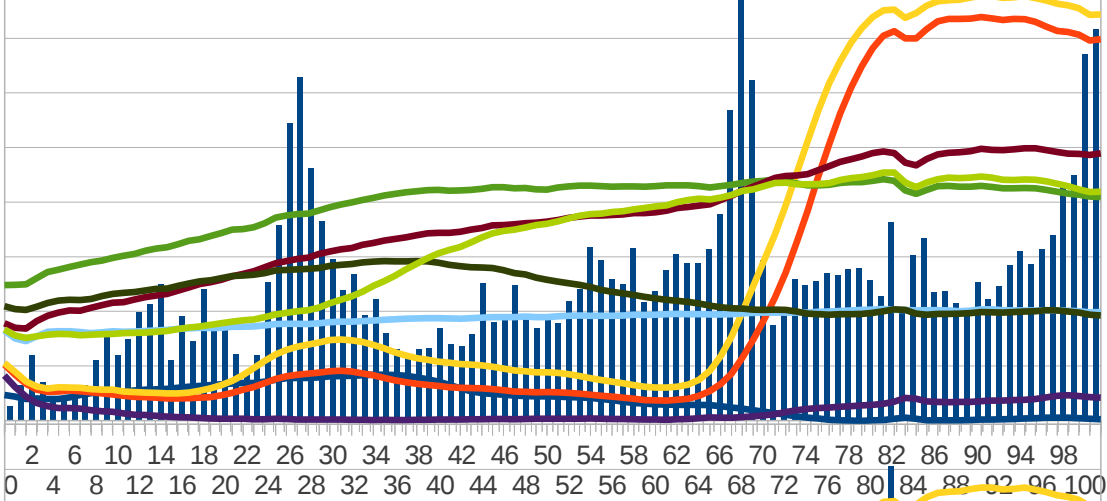
Sélection de 15 bandes



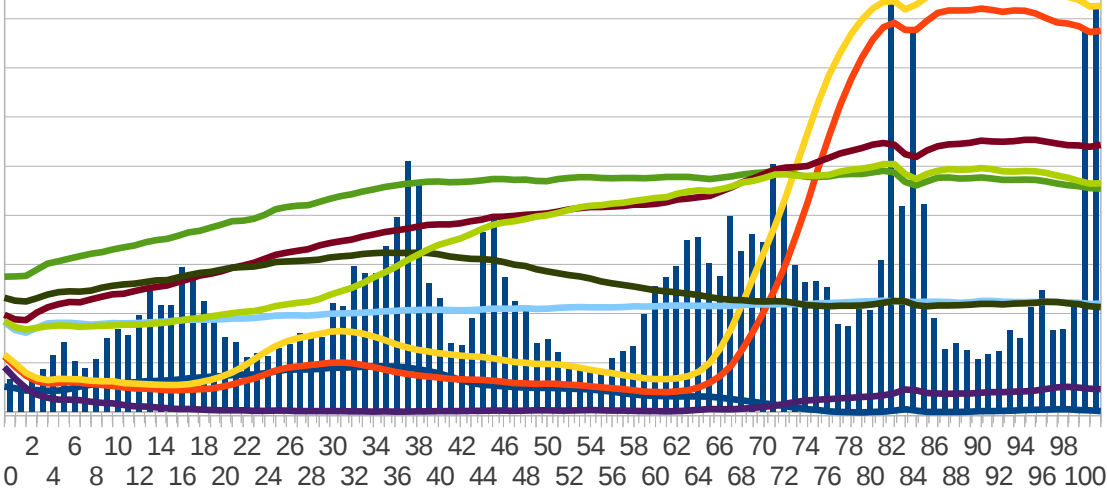
Sélection de 15 bandes



svm linéaire précision moyenne



RF marge



Bdist moyenne



# Jeu test UMBRA Toulouse

Légende encore assez générale :

- Tuile rouge
- Ardoise
- Métal1
- Métal2
  
- Pavés
- Asphalte
- Sol nu
- Voie ferrée
  
- Végétation arborée
- Pelouse

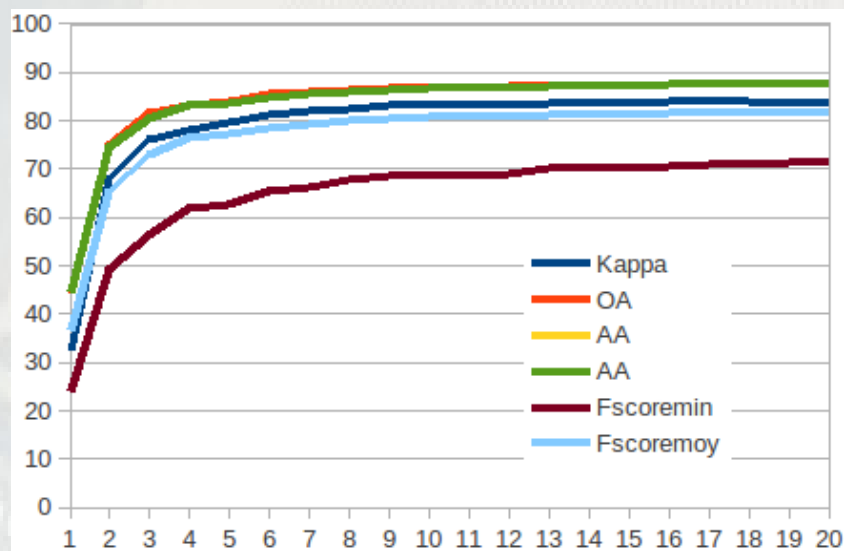
Vérité terrain saisie sur l'une des images acquises à Toulouse (transect )

Données en luminance

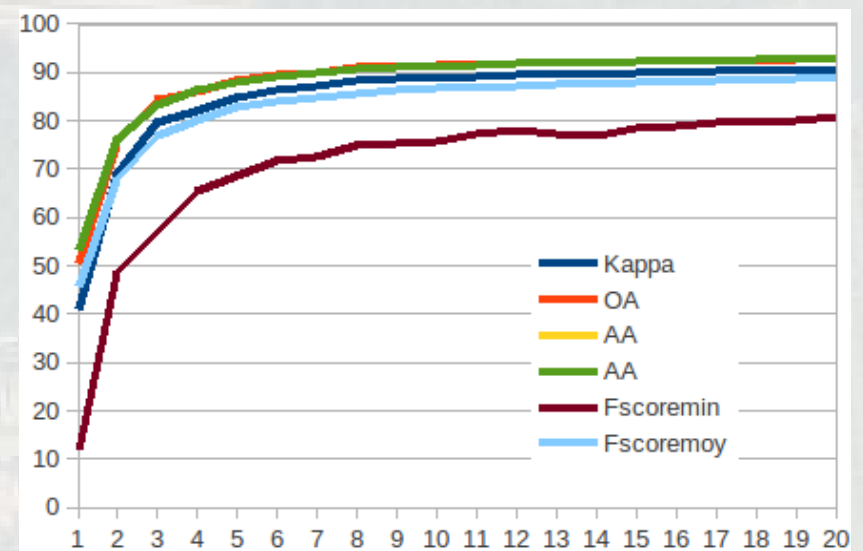
# Jeu test UMBRA

Evolution d'indicateurs de bonne classification en fonction du nombre de bandes sélectionnées par SFFS (score à optimiser = précision moyenne d'un classification SVM linéaire)

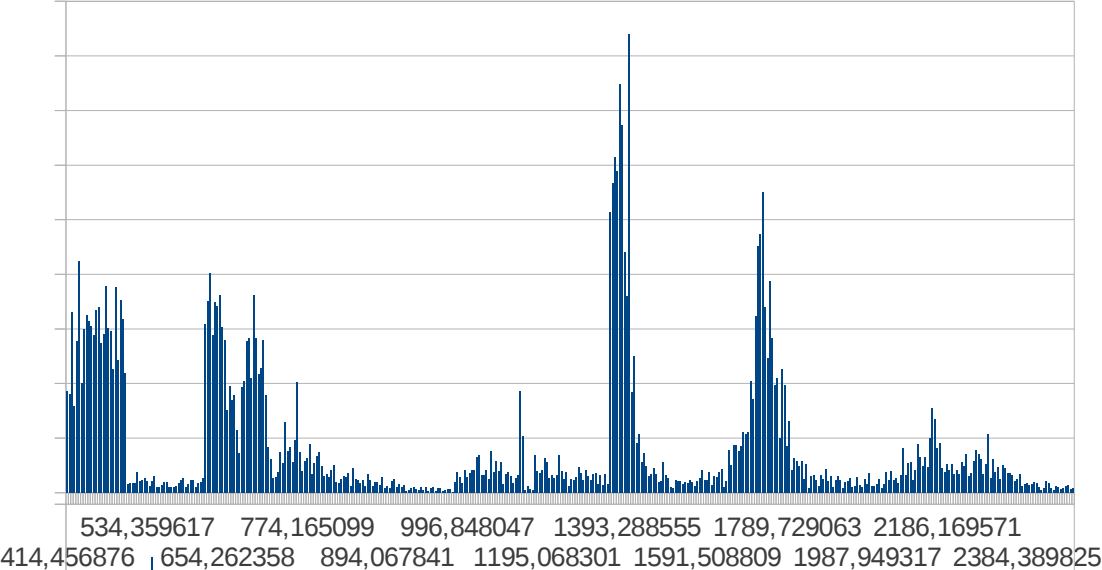
Évaluation par classification RF



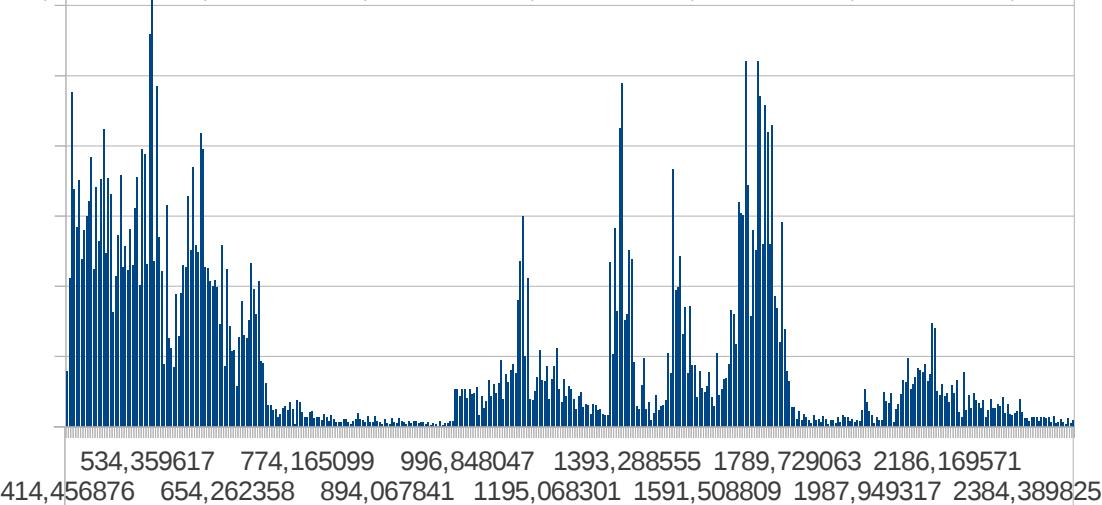
Évaluation par classification SVM gaussienne (avec optimisation)



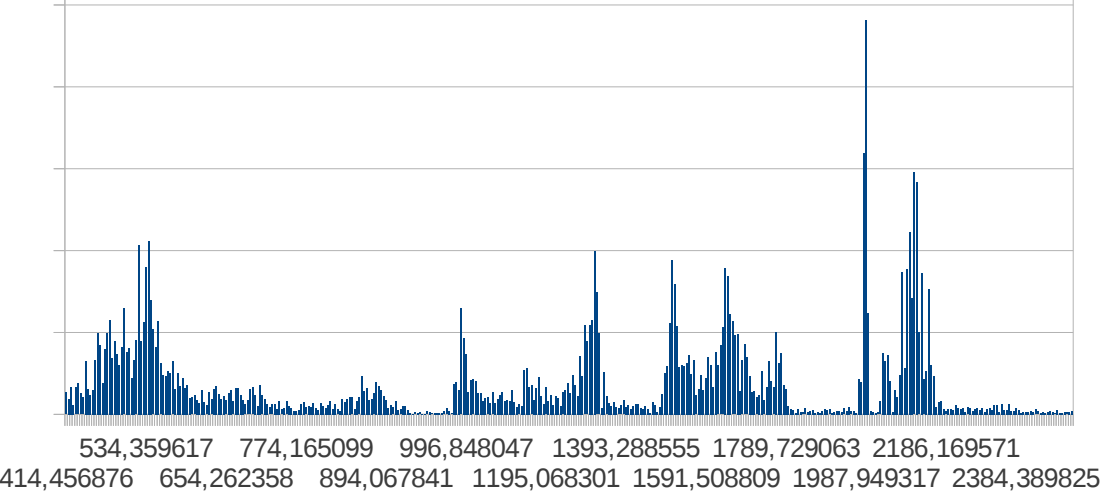
Sélection de 5 bandes



svm linéaire précision moyenne

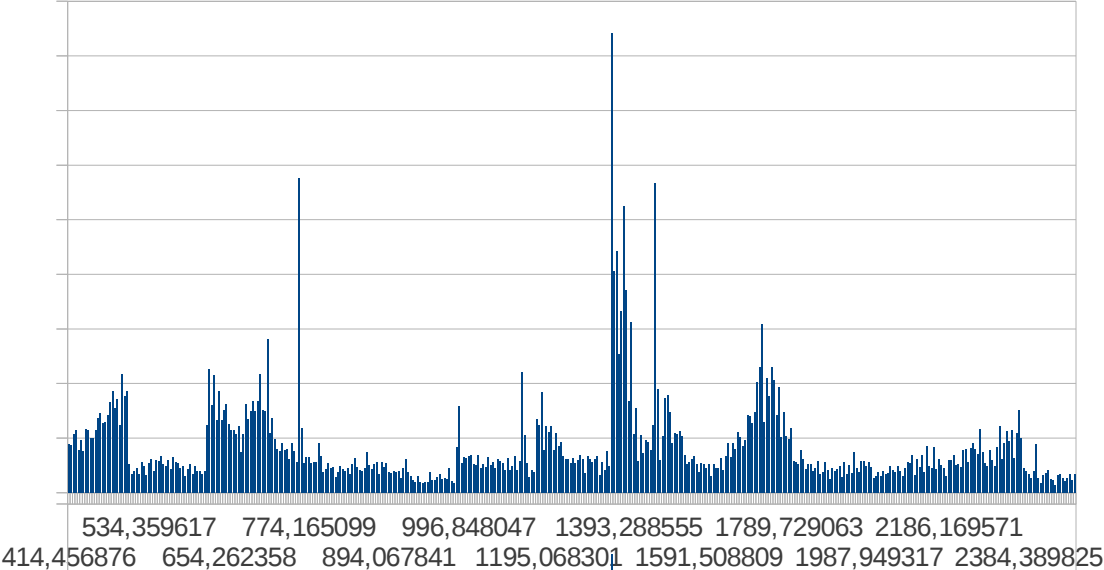


RF marge

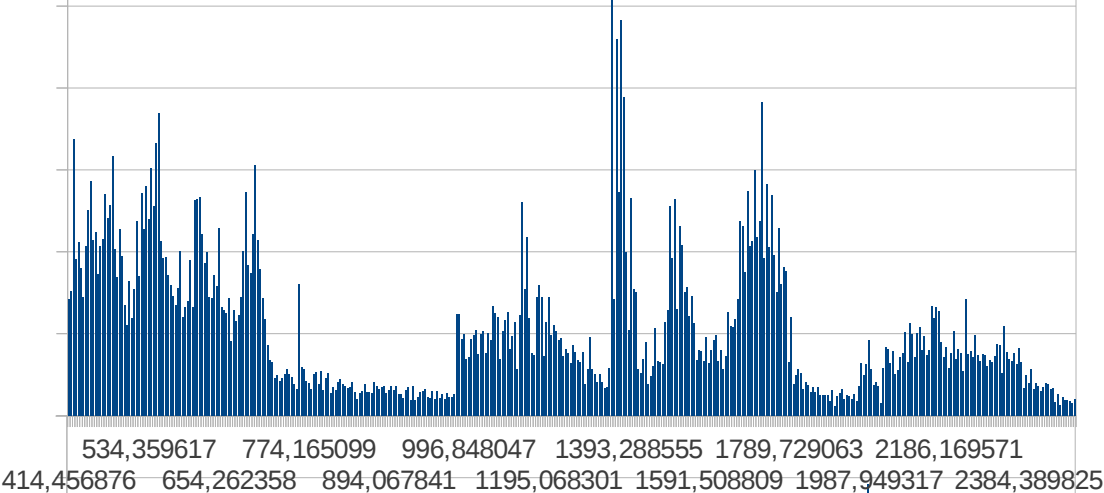


Bdist moyenne

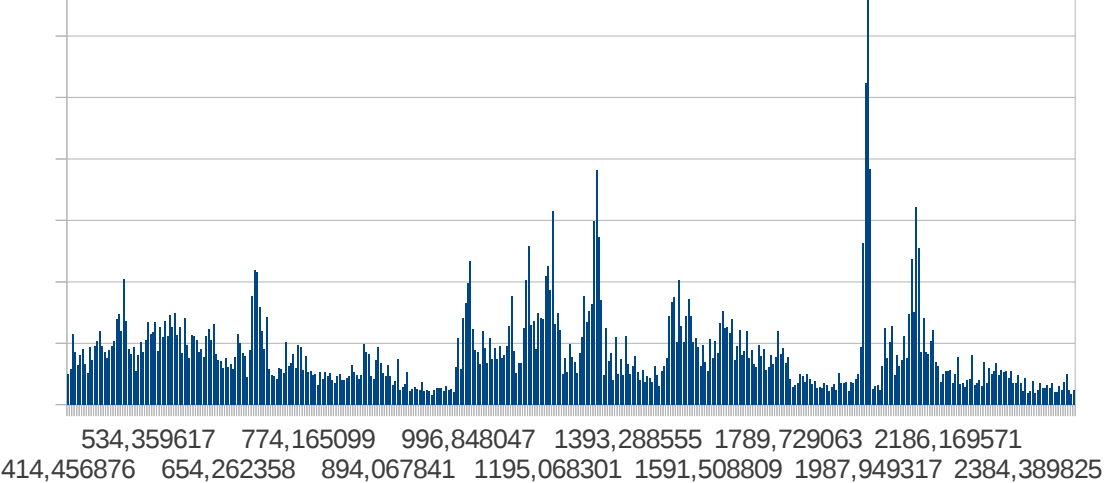
Sélection de 15 bandes



svm linéaire précision moyenne



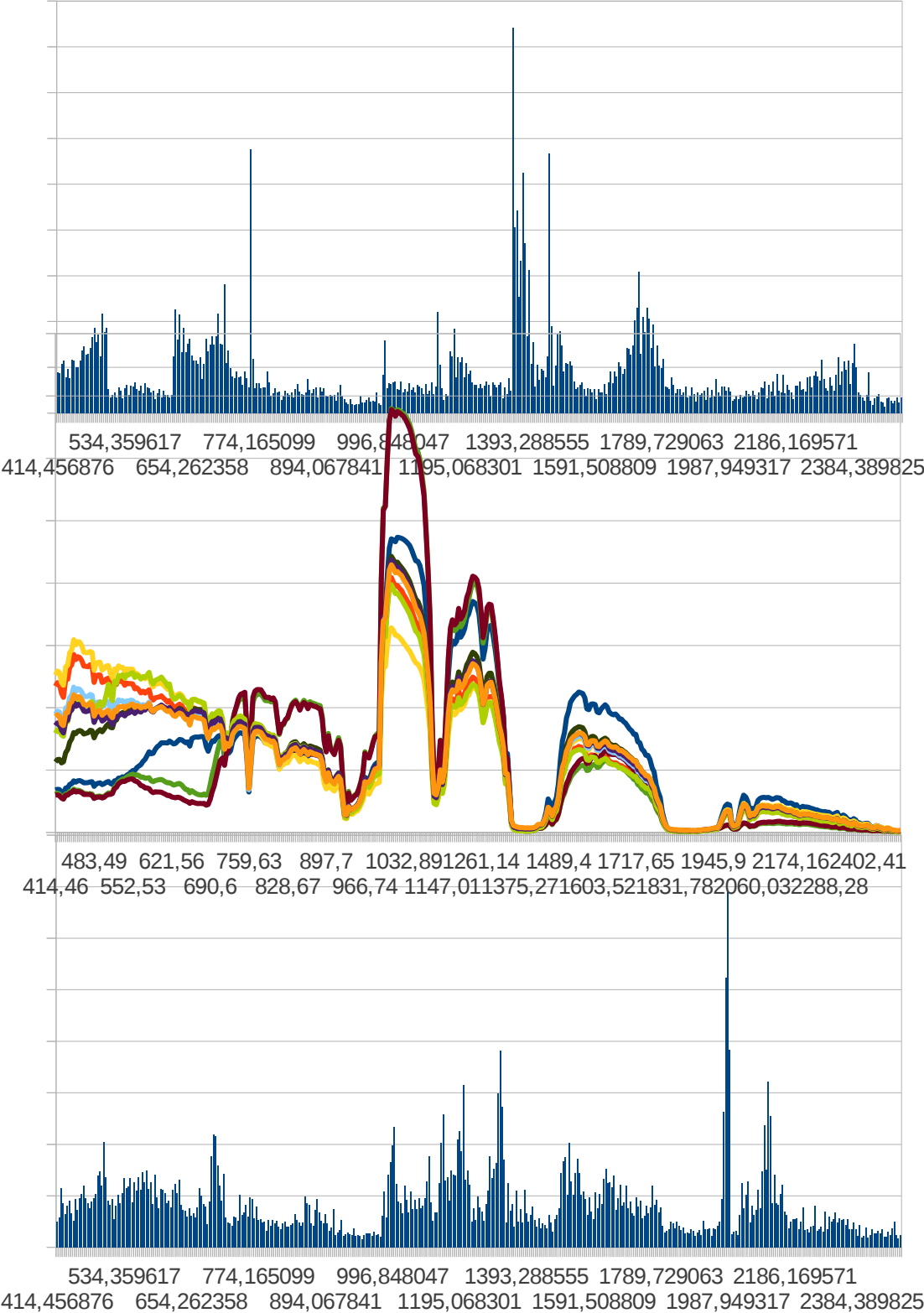
RF marge



Bdist moyenne



Sélection de 15 bandes



svm linéaire précision moyenne

Bdist moyenne

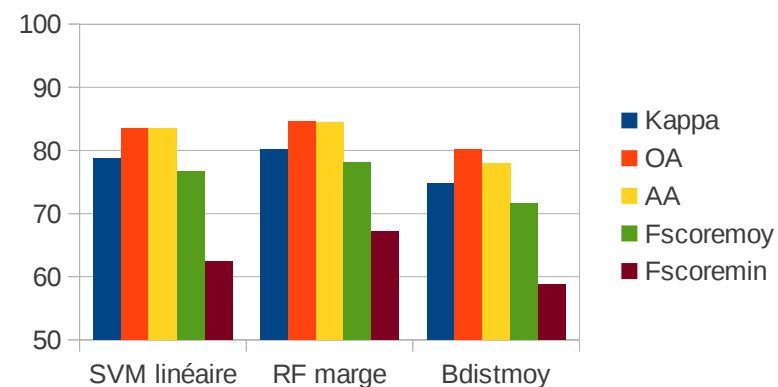
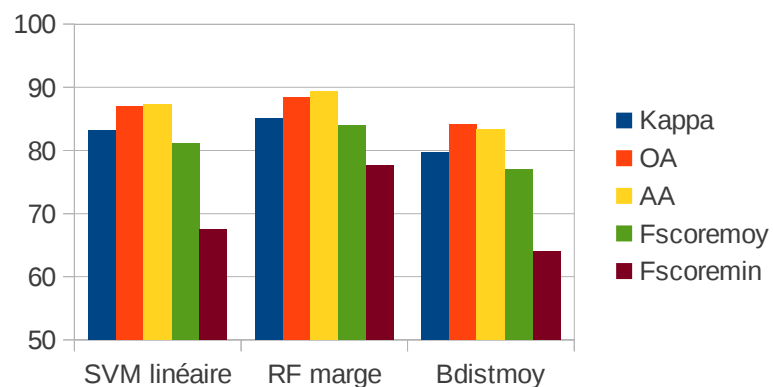
# Jeu test UMBRA

L'un des 3 critères précédents est-il meilleur que les autres ?

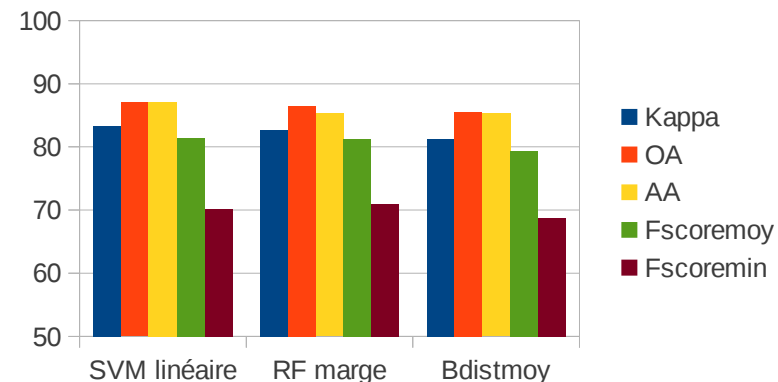
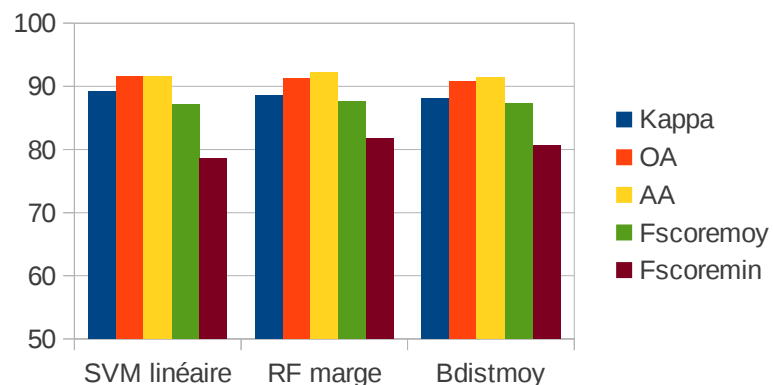
Évaluation par classification SVM gaussienne (avec optimisation)

Évaluation par classification RF

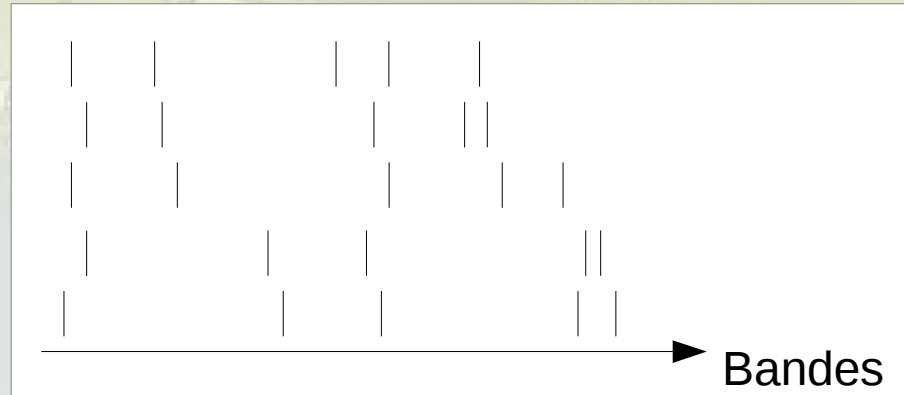
## Sélection de 5 bandes



## Sélection de 15 bandes



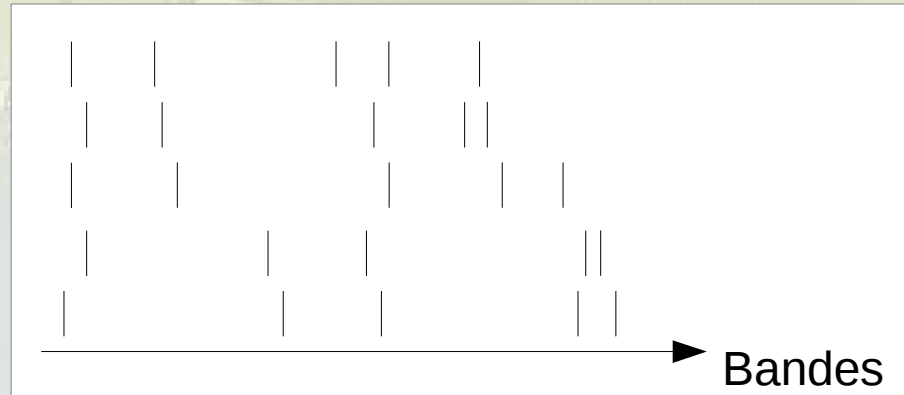
# Perspectives : comportements caractéristiques



Définition d'une distance entre sélections de bandes

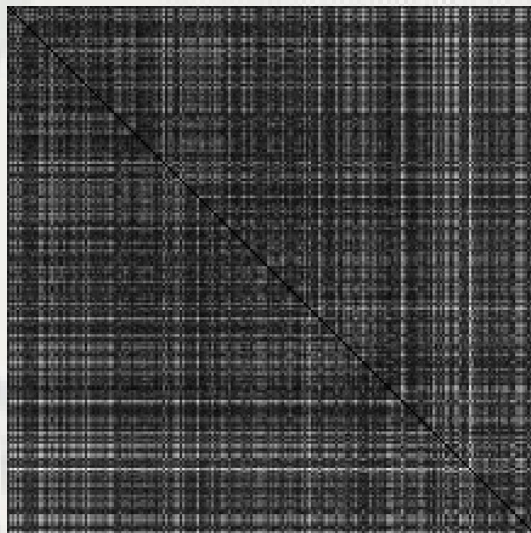


# Perspectives : comportements caractéristiques

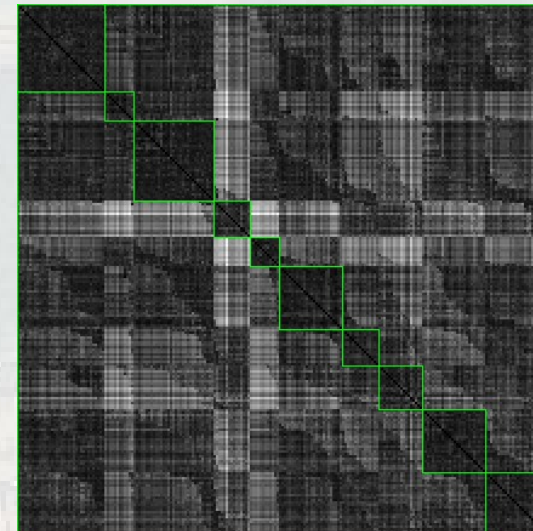


Définition d'une distance entre sélections de bandes

Matrice de distance entre sélections de bandes



clusterisation



# Perspectives : corrélation et fusion

## Regroupement de bandes ayant un comportement semblable

→ Construction de hiérarchies de clusters de bandes

- Fusion hiérarchique des bandes corrélées

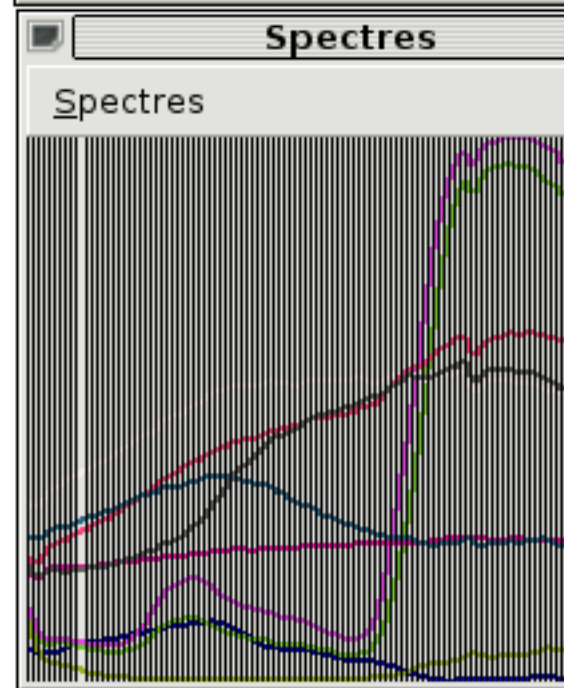
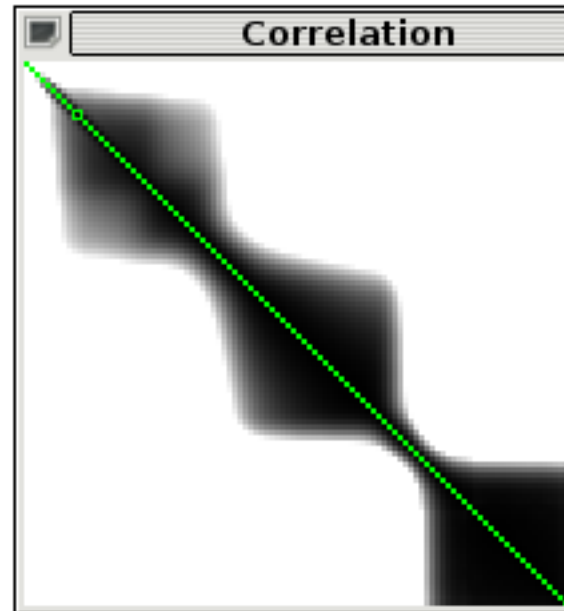
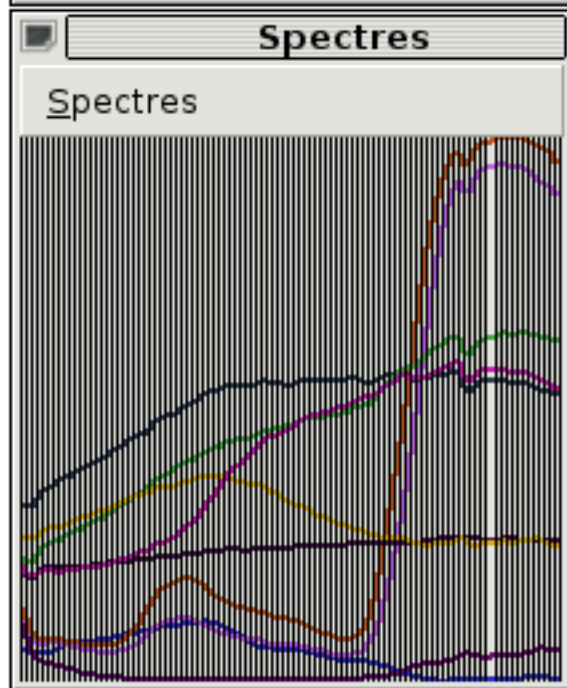
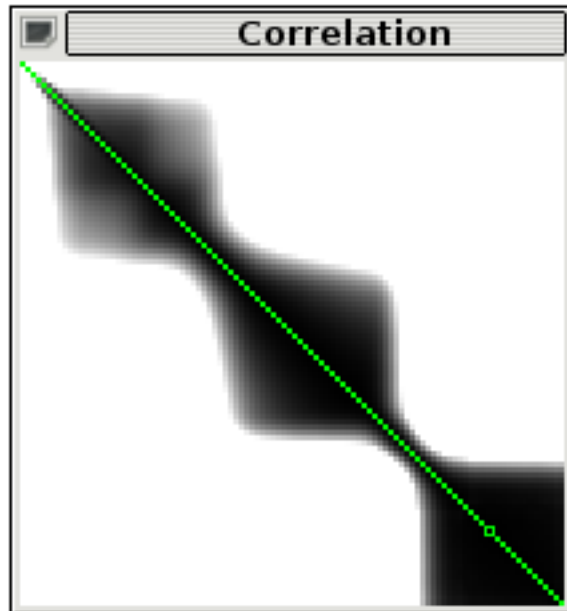
A chaque niveau, on fusionne les clusters de manière à minimiser le score suivant :

$$score(C) = \sum_{c \in C} \sum_{i \in c, j \in c, i \neq j} 1 - \text{correl}(i, j)$$

- Fusion hiérarchique des bandes en fonction de leur contribution à la séparabilité

A chaque niveau, on fusionne les 2 clusters dont la fusion diminue le moins la séparabilité.

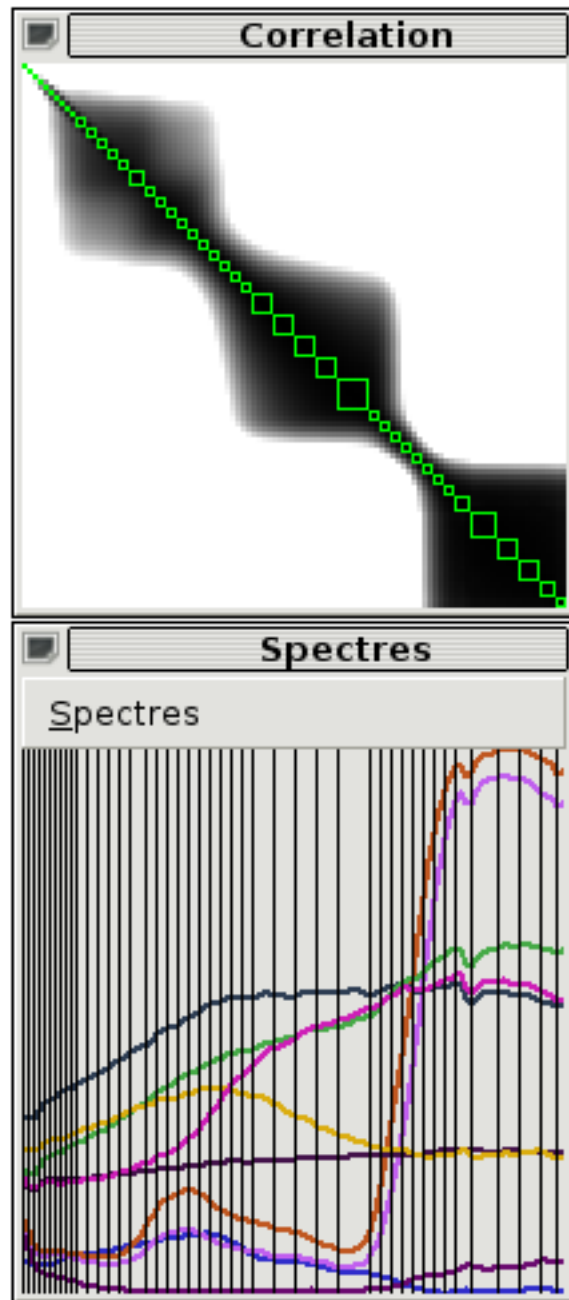
# Perspectives : corrélation et fusion



Regroupement  
basé sur la  
corrélation

Regroupement  
basé sur la  
séparabilité

# Perspectives : corrélation et fusion



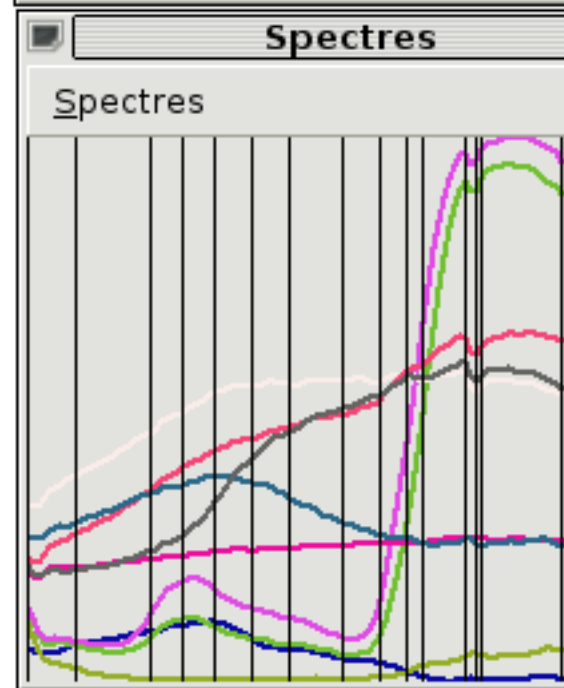
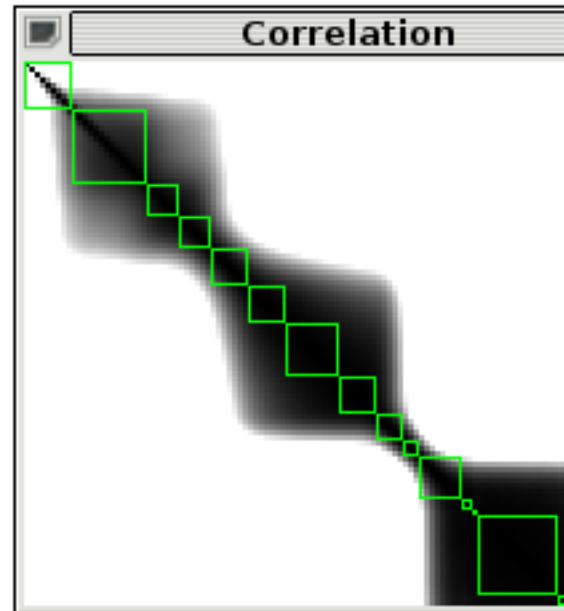
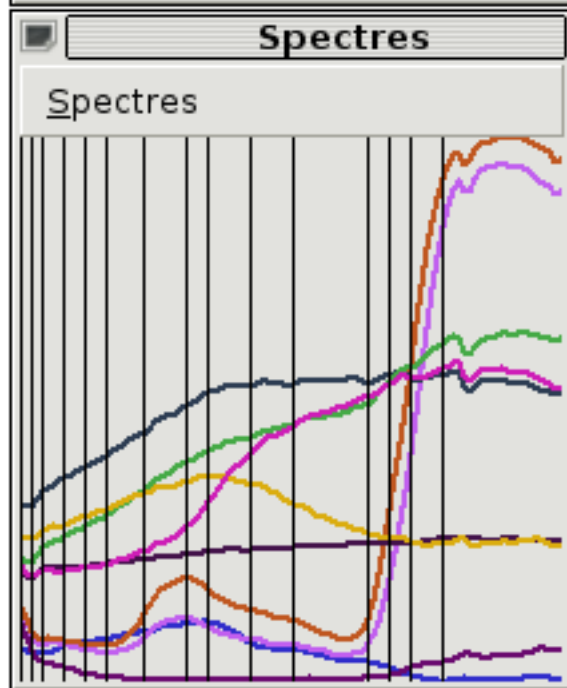
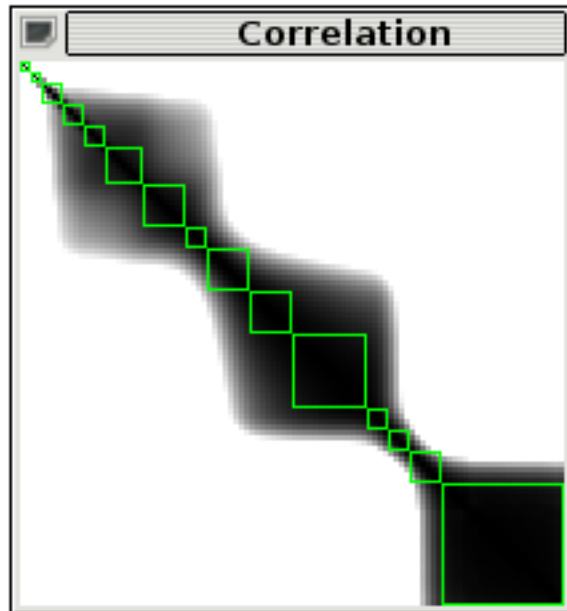
Regroupement  
basé sur la  
corrélation



Regroupement  
basé sur la  
séparabilité



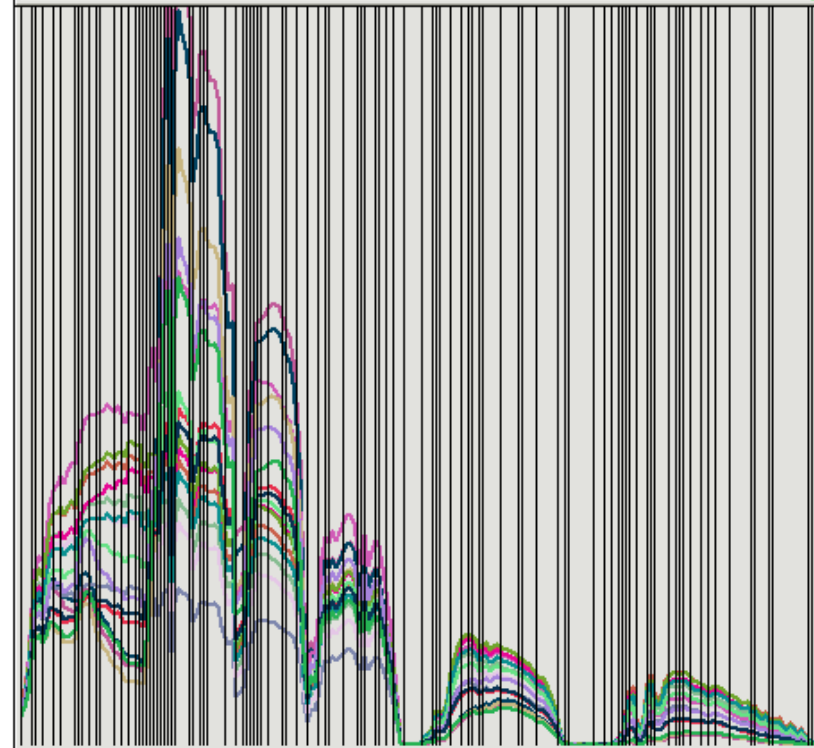
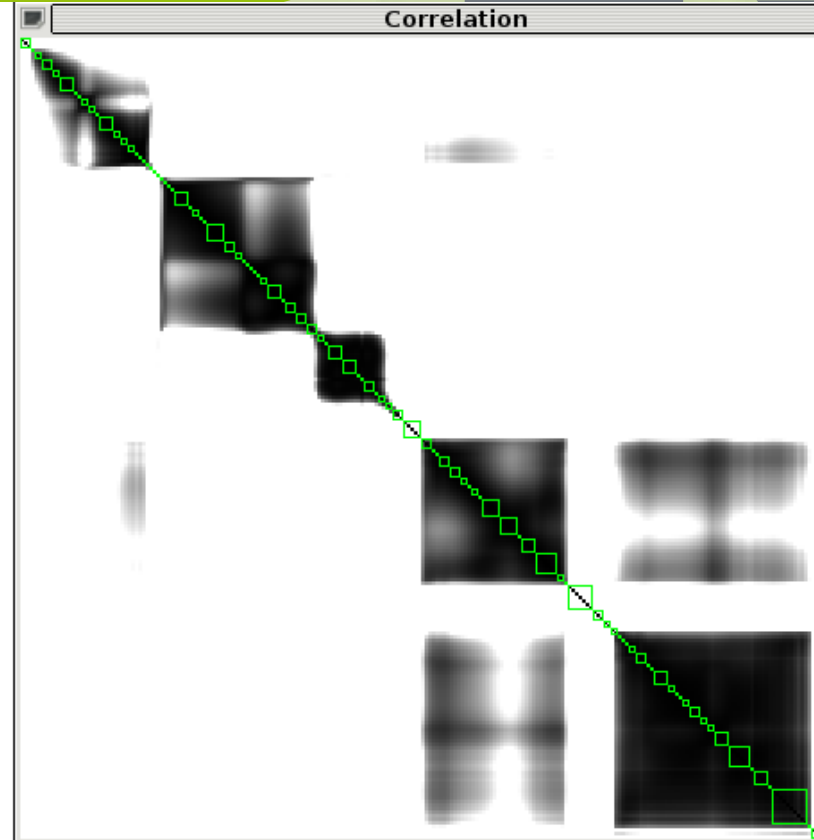
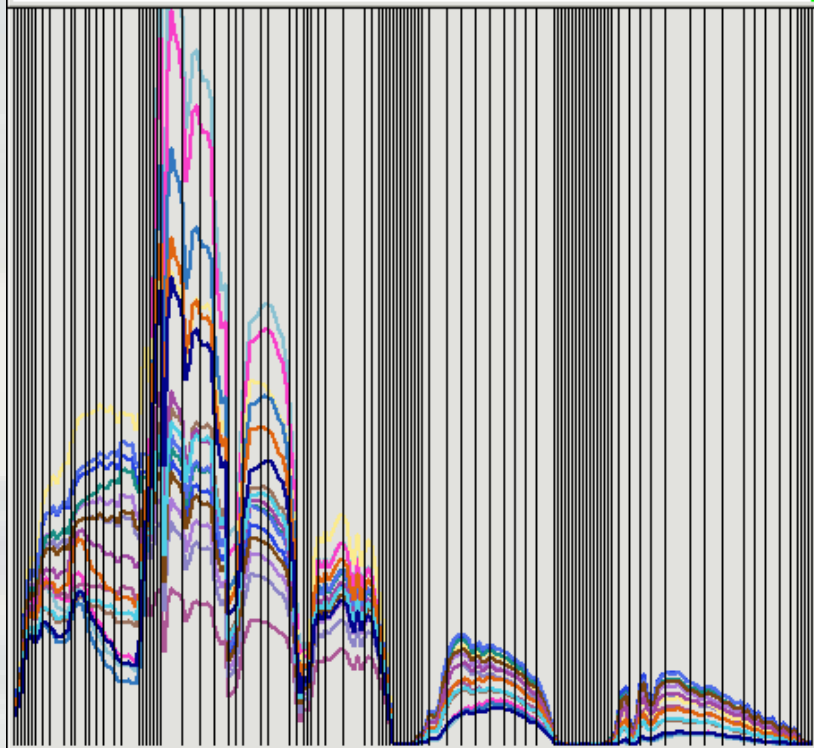
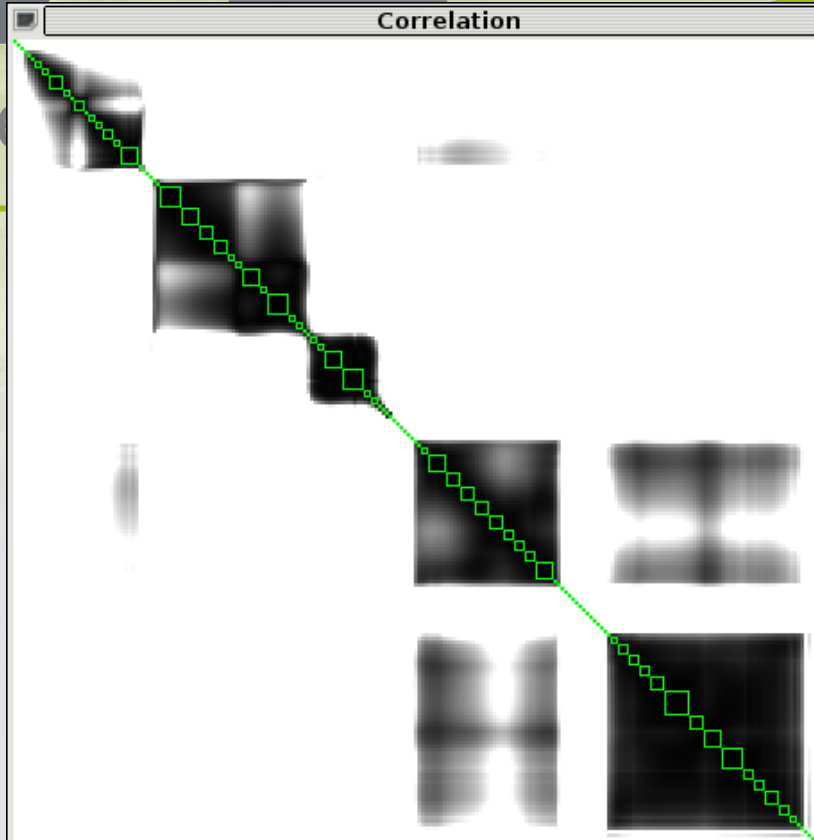
# Perspectives : corrélation et fusion



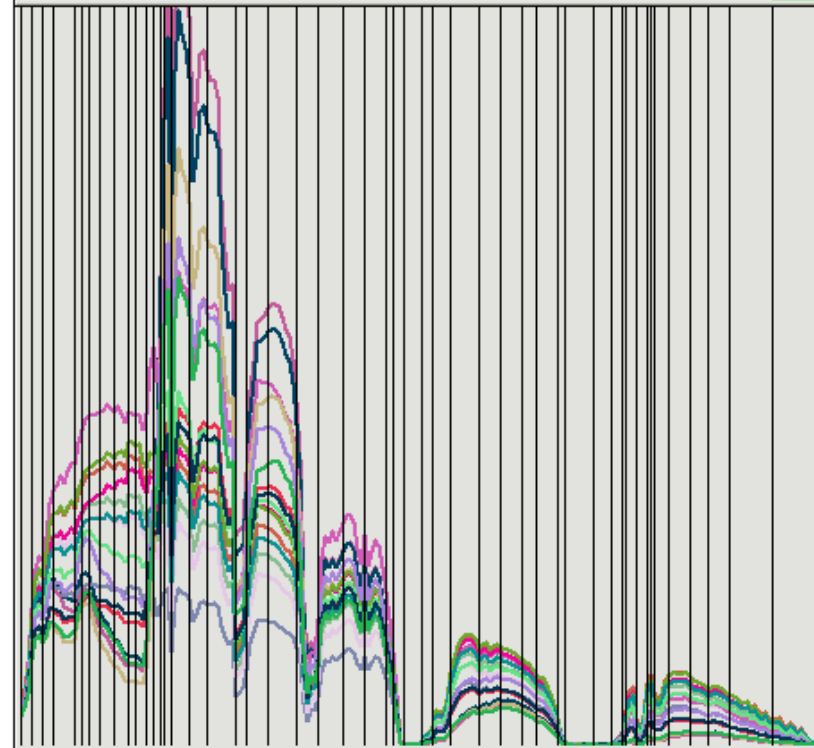
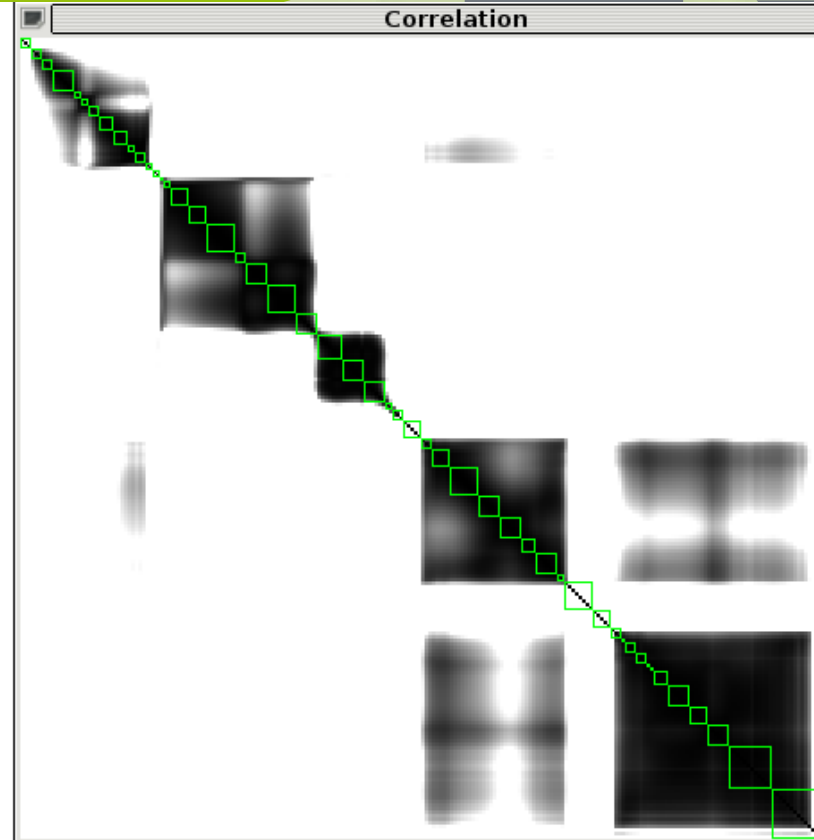
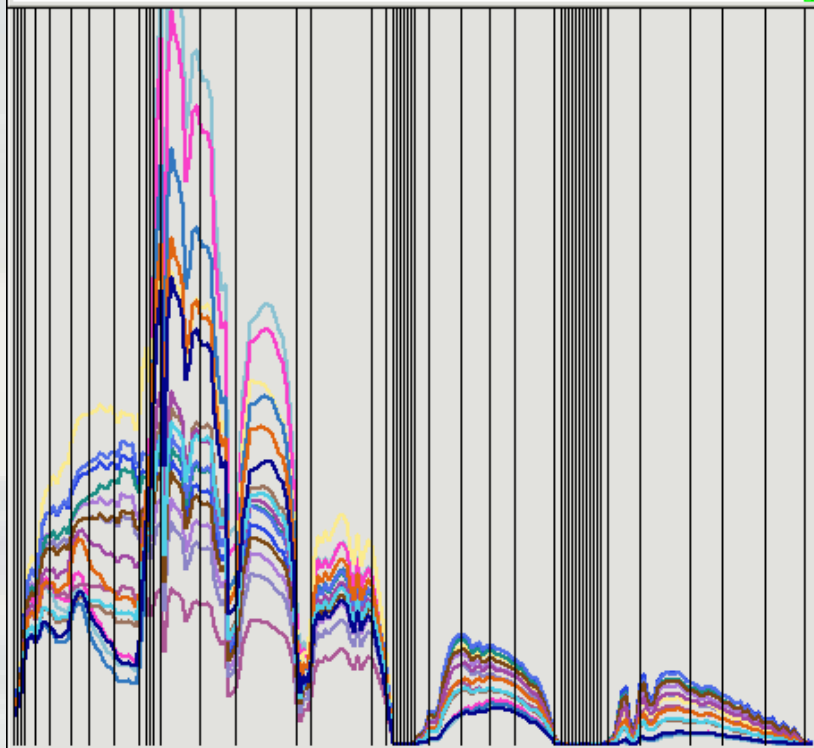
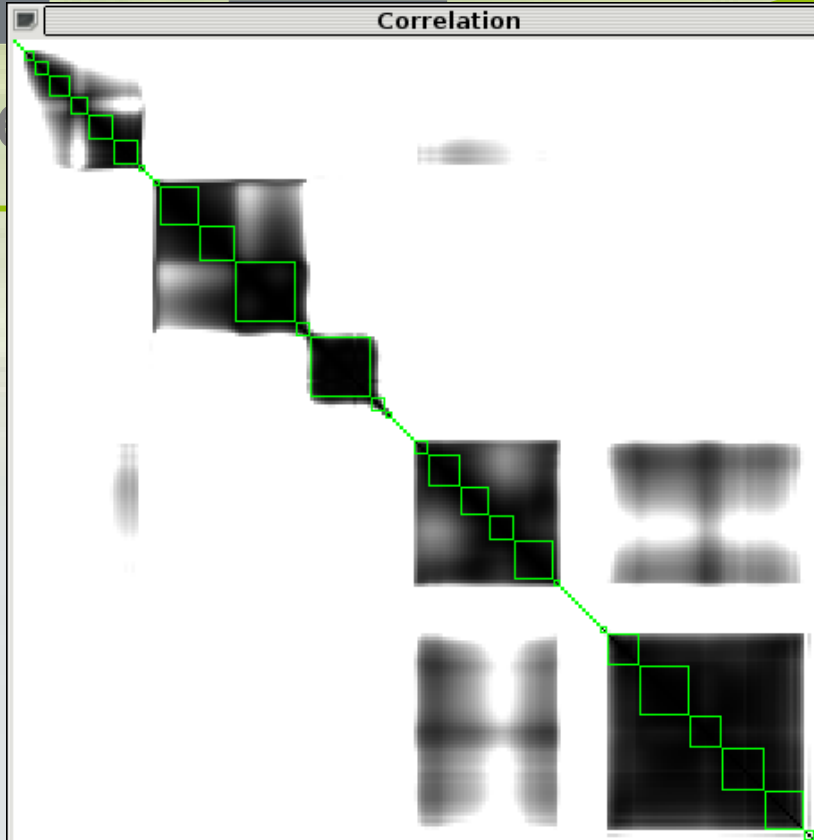
Regroupement  
basé sur la  
corrélation

Regroupement  
basé sur la  
séparabilité

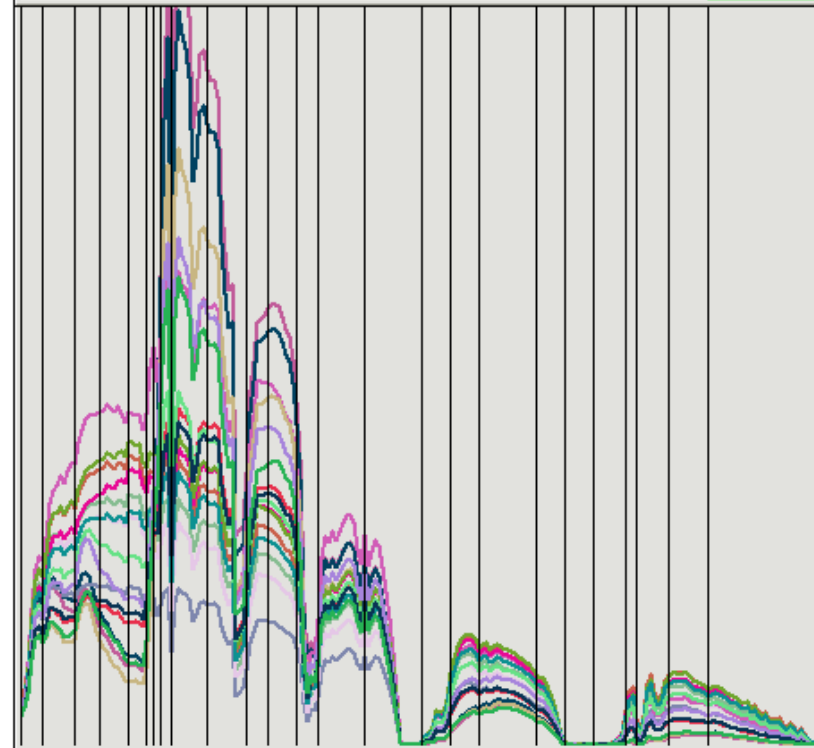
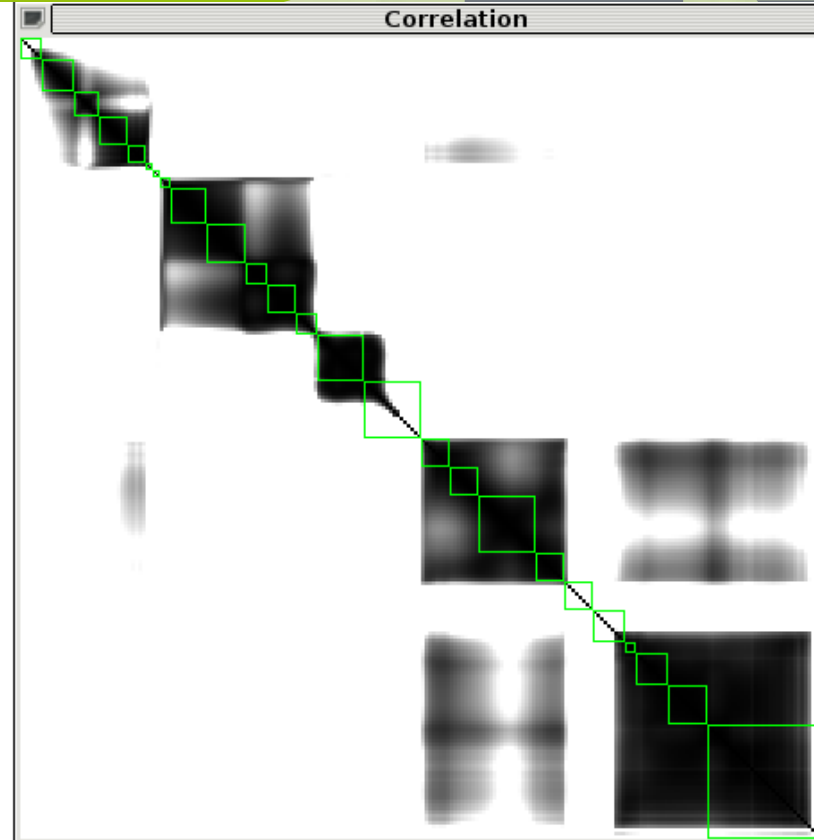
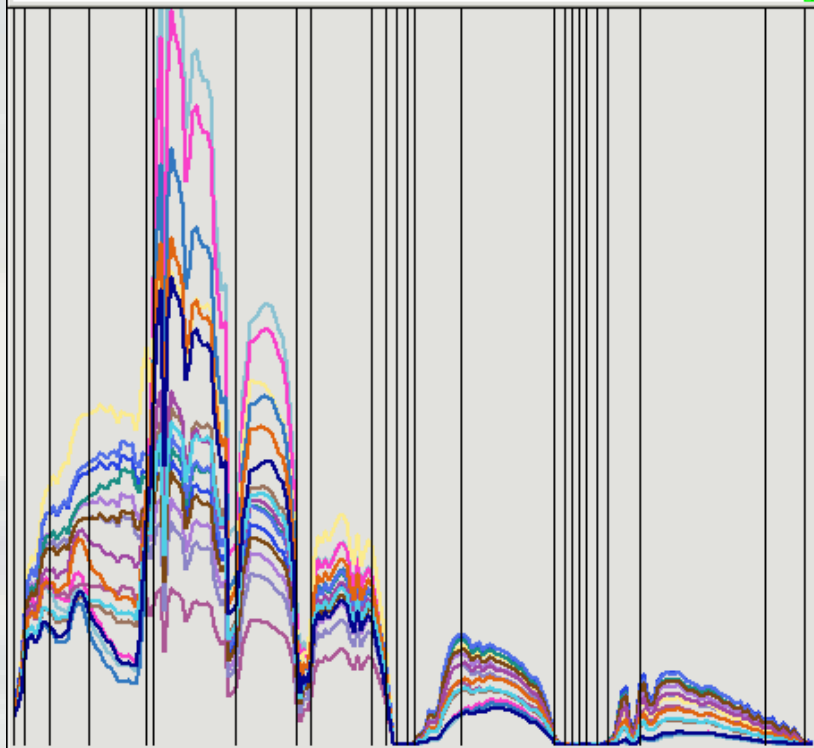
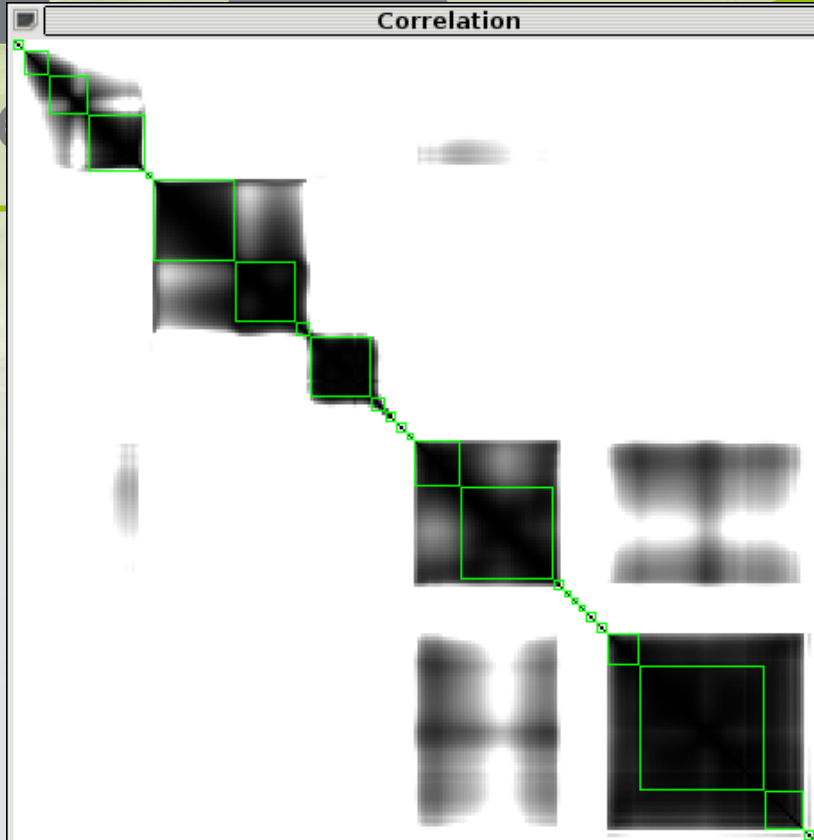
PO



PO



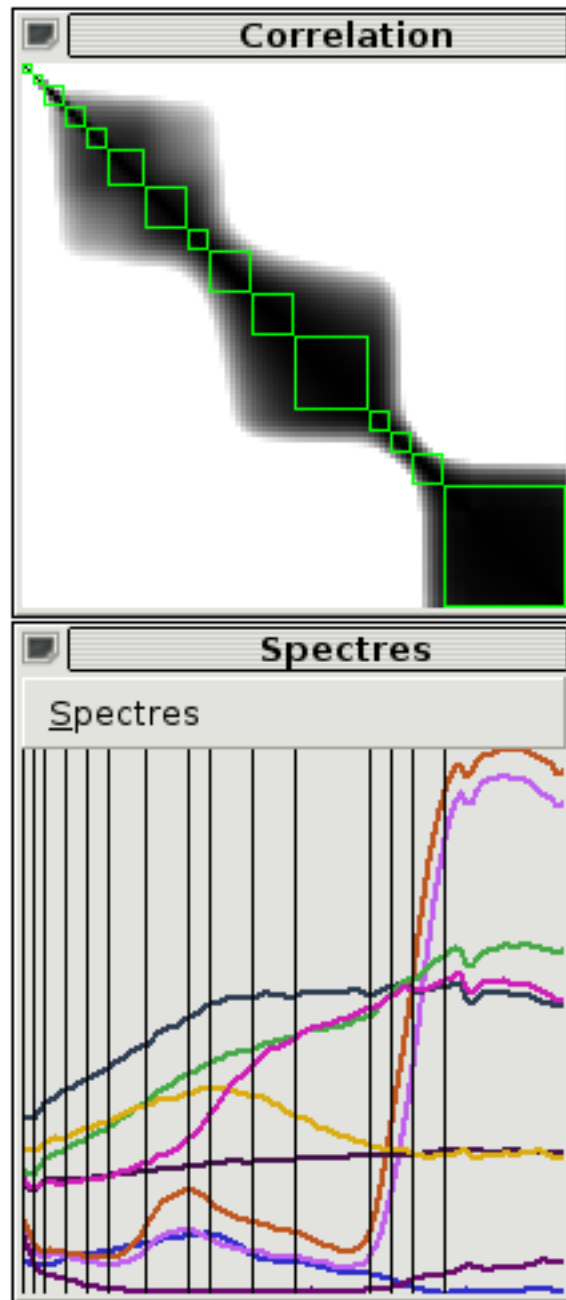
PO



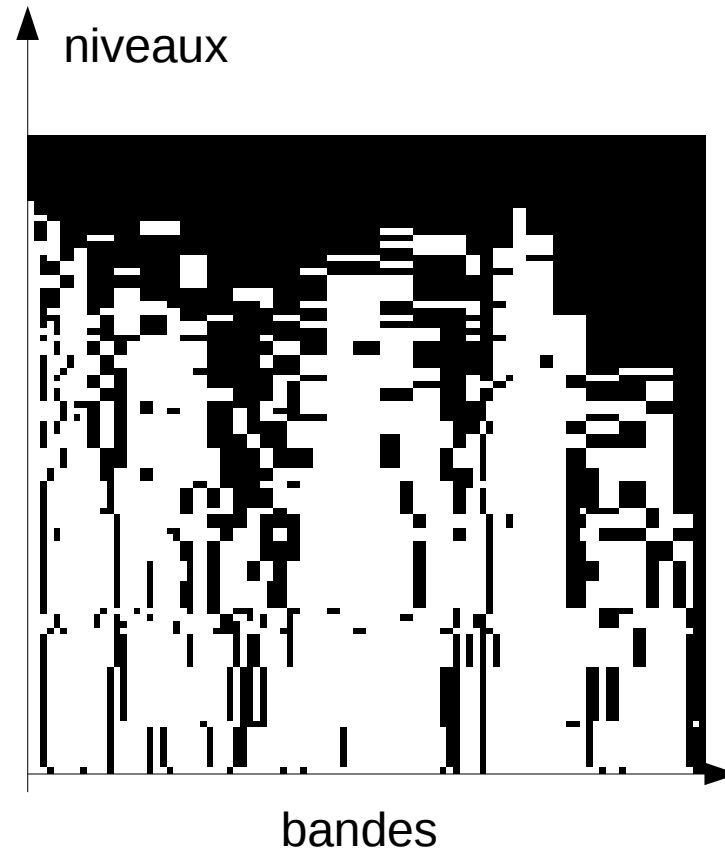


# Perspectives : corrélation et fusion

Regroupement  
 basé sur la  
 corrélation



Ex : sélection de 15 bandes  
 à chaque niveau de la hiérarchie





Merci pour votre attention.