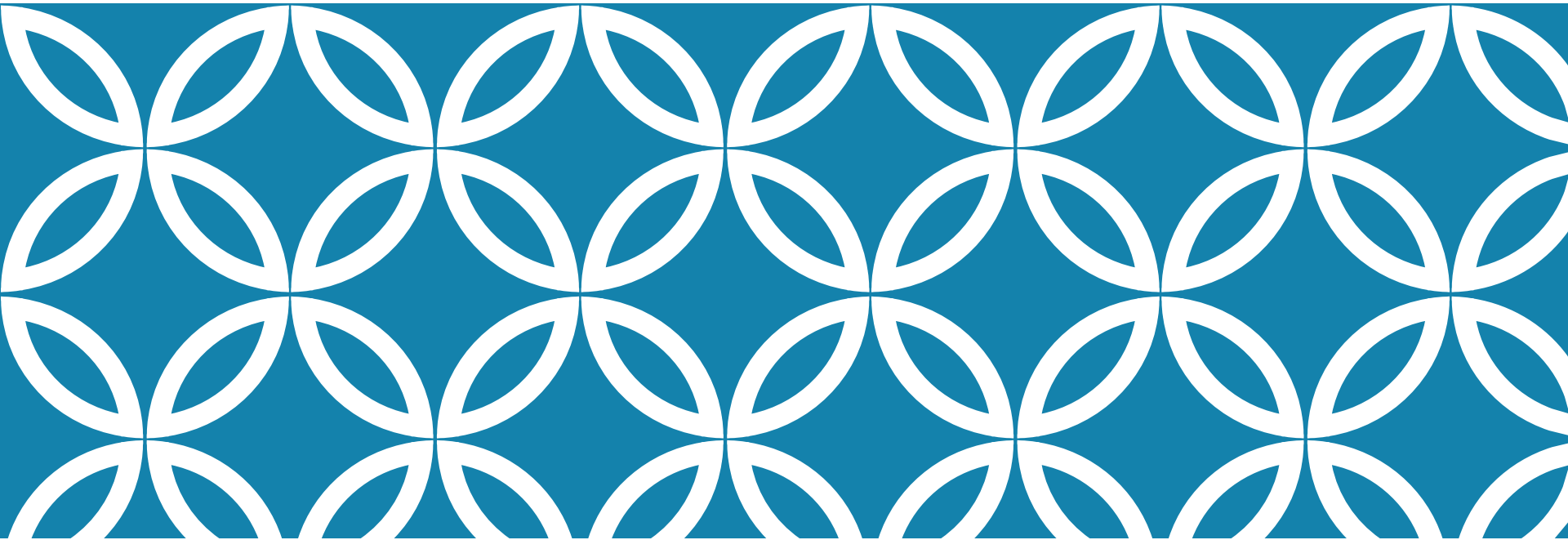




Présentation à TerraNIS



# PROJET OENOMIP: VOLET HYPERSPECTRAL

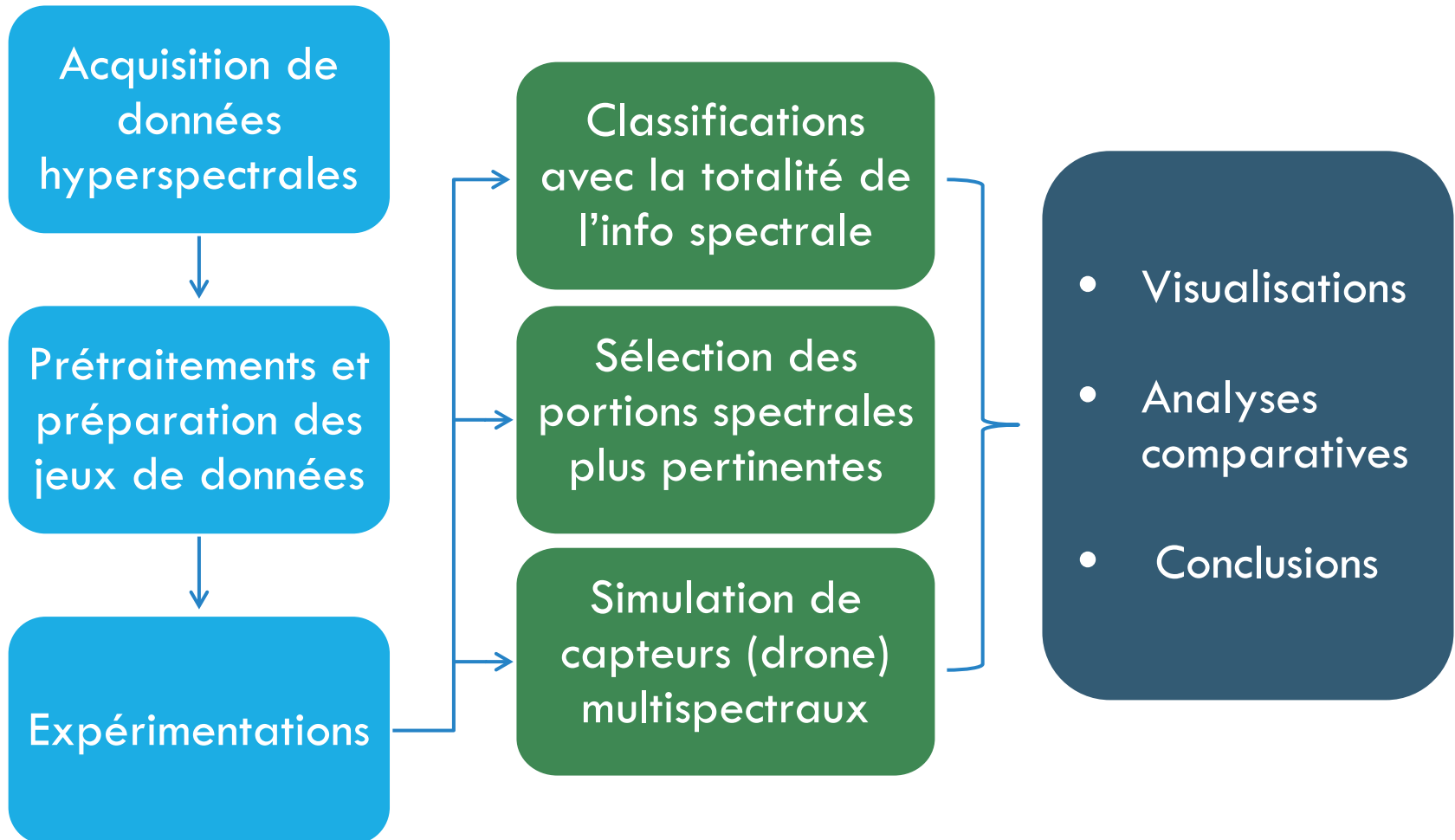
Aperçu méthodologique et  
des résultats obtenus

*Post-doc : Fabio Guttler*

*Responsables scientifiques: Sylvie Duthoit, Mathieu  
Fauvel et Anne Jacquin*

Toulouse, 14 Avril 2017

# SCHÉMA MÉTHODOLOGIQUE GÉNÉRAL



# ACQUISITION DONNÉES

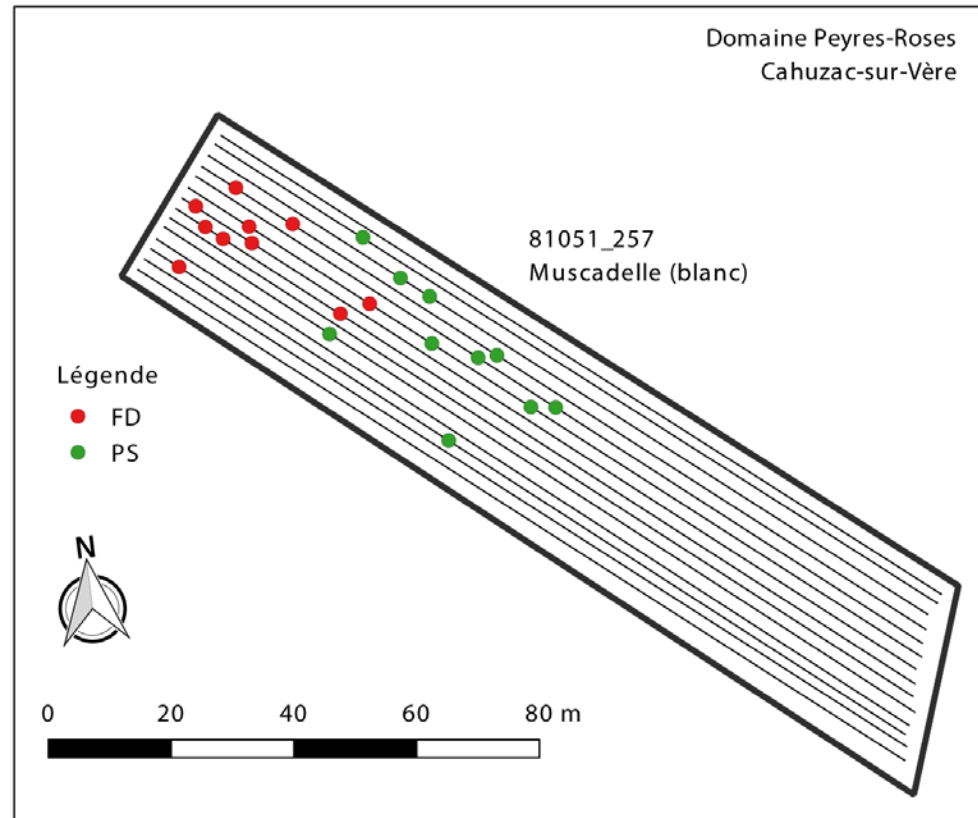
## Protocole de terrain

Par parcelle de vigne :

- 10 pieds sains (PS)
- 10 pieds infectés par la Flavescence dorée (FD)

\* Uniquement pour la campagne 2016 :

- 10 pieds infectés par des maladies du bois (Esca, BDA)



# ACQUISITION DONNÉES

## Aperçu des décolorations foliaires

Feuille saine

Flavescence dorée

Esca



Cépage :  
Cabernet Sauvignon

## Protocole de terrain

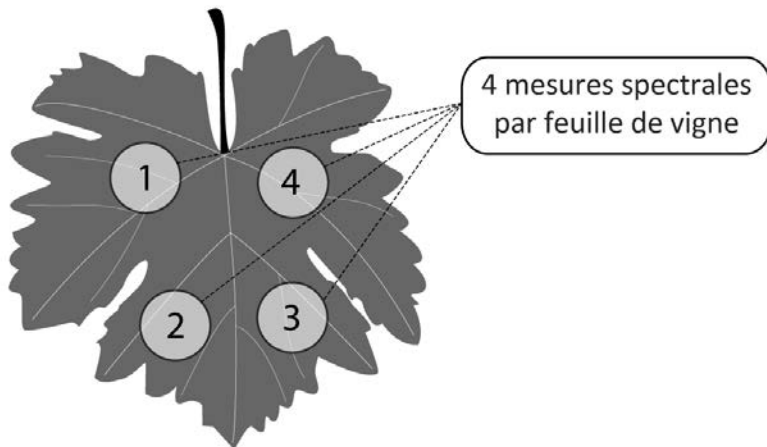
- Prélèvement de 4 feuilles par pied de vigne



Cépage :  
Sirah

# ACQUISITION DONNÉES

*Spectroradiomètre*



## Protocole de terrain

Prélèvement de 4 feuilles par pied de vigne avec 4 mesures par feuille : au total, 16 mesures par pied.



# ACQUISITION DONNÉES

## Données récoltées

Sur un total de :

- 17 parcelles de vigne
- 8 cépages rouges
- 6 cépages blancs

Année	Couleur	Cépage	Site
2015	Rouge	Braucol	Gaillacois
		Duras	
		Gamay (2 parcelles)	
	Blanc	Colombard	
		Loin de l'Œil	
		Muscadelle	
		Sauvignon	
2016	Rouge	Braucol	Gaillacois
		Duras	
		Merlot	
		Cabernet-Sauvignon	Minervois
		Grenache	
		Mourvèdre	
		Syrah	
	Blanc	Ondenc	Gaillacois
		Muscadelle	
		Chardonnay	Minervois

# PRÉTRAITEMENTS

**Jeu de données  
« brut »**

Année	Nombre de pieds	Nombre de mesures
2015	168	2 688
2016	280	4 469

# PRÉTRAITEMENTS

## Jeu de données « brut »

Année	Nombre de pieds	Nombre de mesures
2015	168	2 688
2016	280	4 469

Etat de santé	Nombre de pieds	Nombre de mesures
Flavescence dorée	181	2 892
Sain	180	2 880
Esca	87	1 385

Soit un total de 7 157 mesures spectrales :  
Ce qui représente 7 157 fichiers .asd (35 ko chacun => env. 250 Mo)

### 1<sup>er</sup> prétraitement :

- exporter fichiers au format .ascii (56 ko chacun => env. 400 Mo)



# PRÉTRAITEMENTS

## Correction des sauts spectraux

Sauts entre les trois capteurs du spectro :  
VNIR – SWIR1 – SWIR2

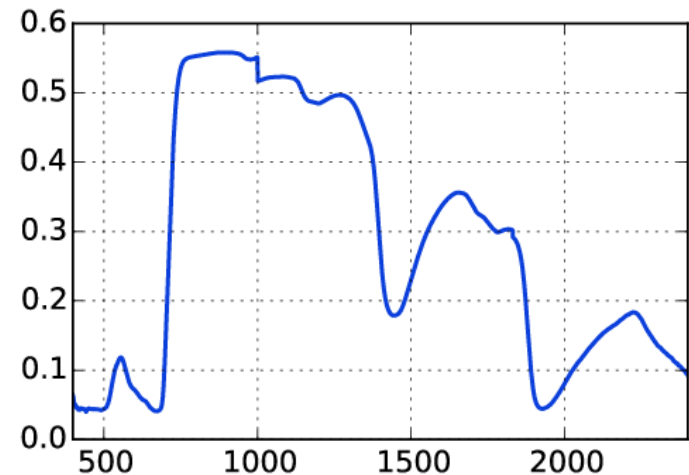
c'est-à-dire entre les longueurs d'onde :

- 1000 et 1001 nm
- 1800 et 1801 nm

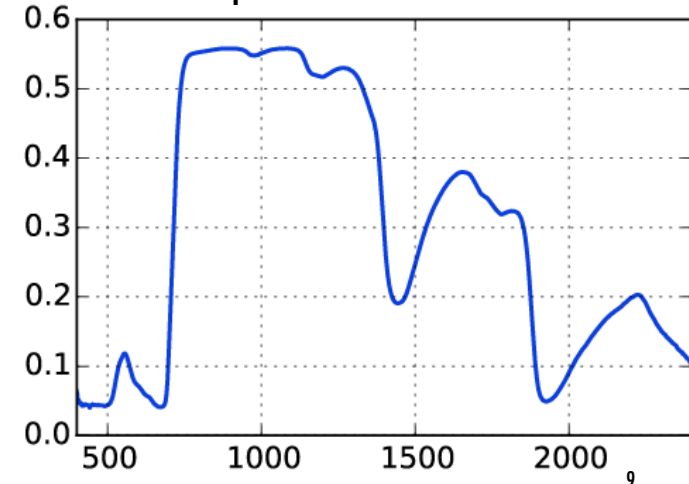
Correction par la méthode de la tangente,  
supposant un factor multiplicatif entre les  
valeurs perturbées et les valeurs attendues

Calcul de la dérivée à 1000, 1001,  
1800 et 1801

Avant correction



Après correction



# PRÉTRAITEMENTS

## Suppression portions bruitées du spectre

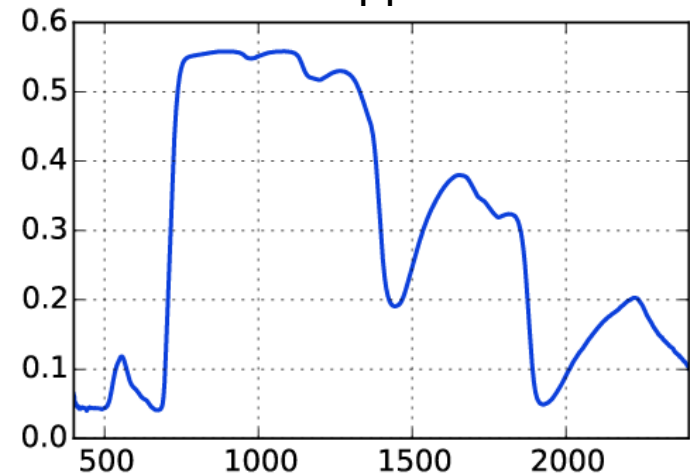
- $< 400$  nm
- $> 2400$  nm

## Suppression bandes d'absorption d'eau

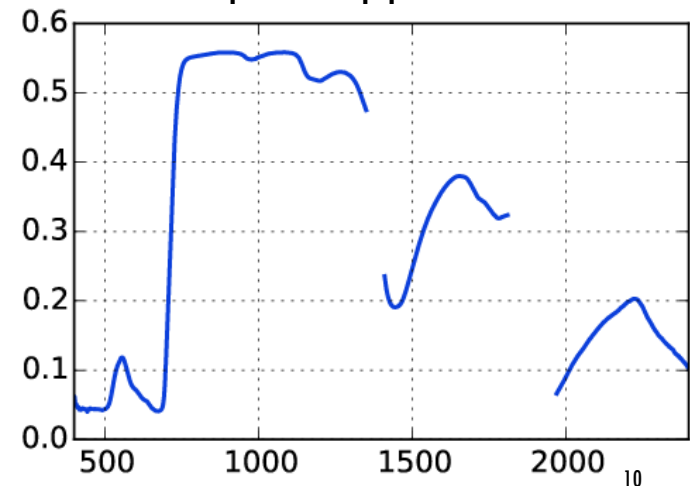
- entre 1350 et 1409 nm
- entre 1810 et 1969 nm

De ce fait, des 2200 longueurs d'onde initiales, 1783 en sont conservées.

Avant suppression



Après suppression

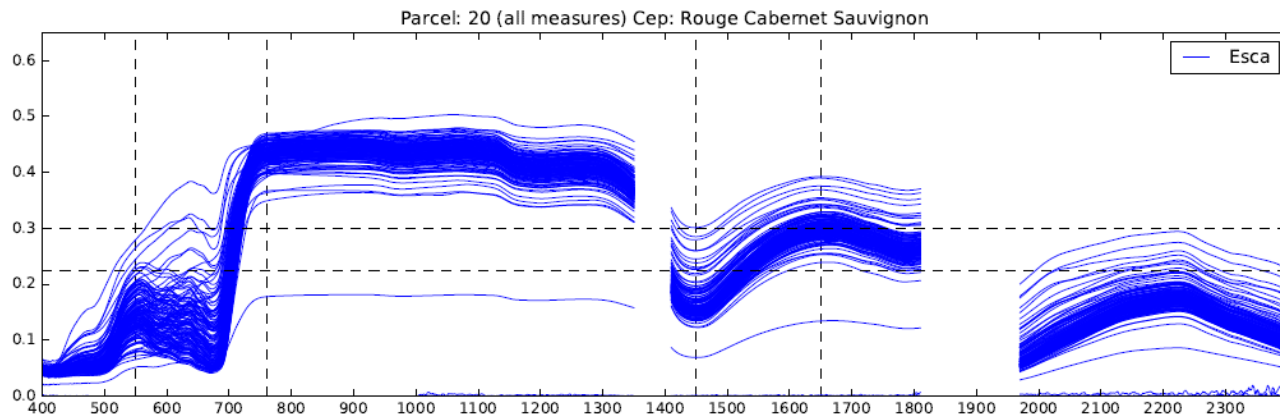


# PRÉTRAITEMENTS

## Détection de spectres aberrants

4 types d'erreurs :

- a) Spectres possédant des valeurs  $> 1$
- b) spectres aux valeurs nulles** (proche de 0 sur toutes les longueurs d'onde)
- c) spectres présentant des valeurs de réflectance exceptionnellement faibles (voire négatives) dans l'intervalle 400-500 nm
- d) spectres présentant une forme à priori normale mais avec des valeurs de réflectance étrangement faibles**

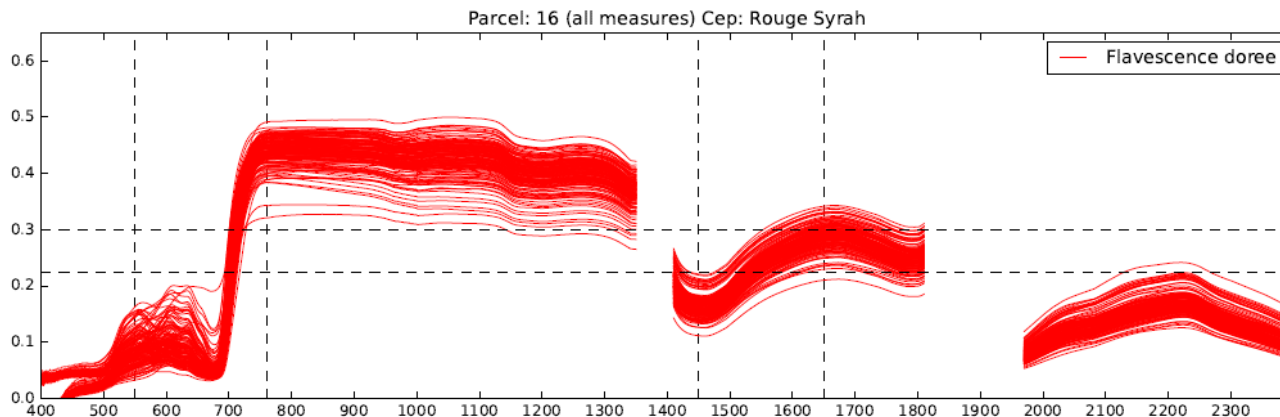


# PRÉTRAITEMENTS

## Détection de spectres aberrants

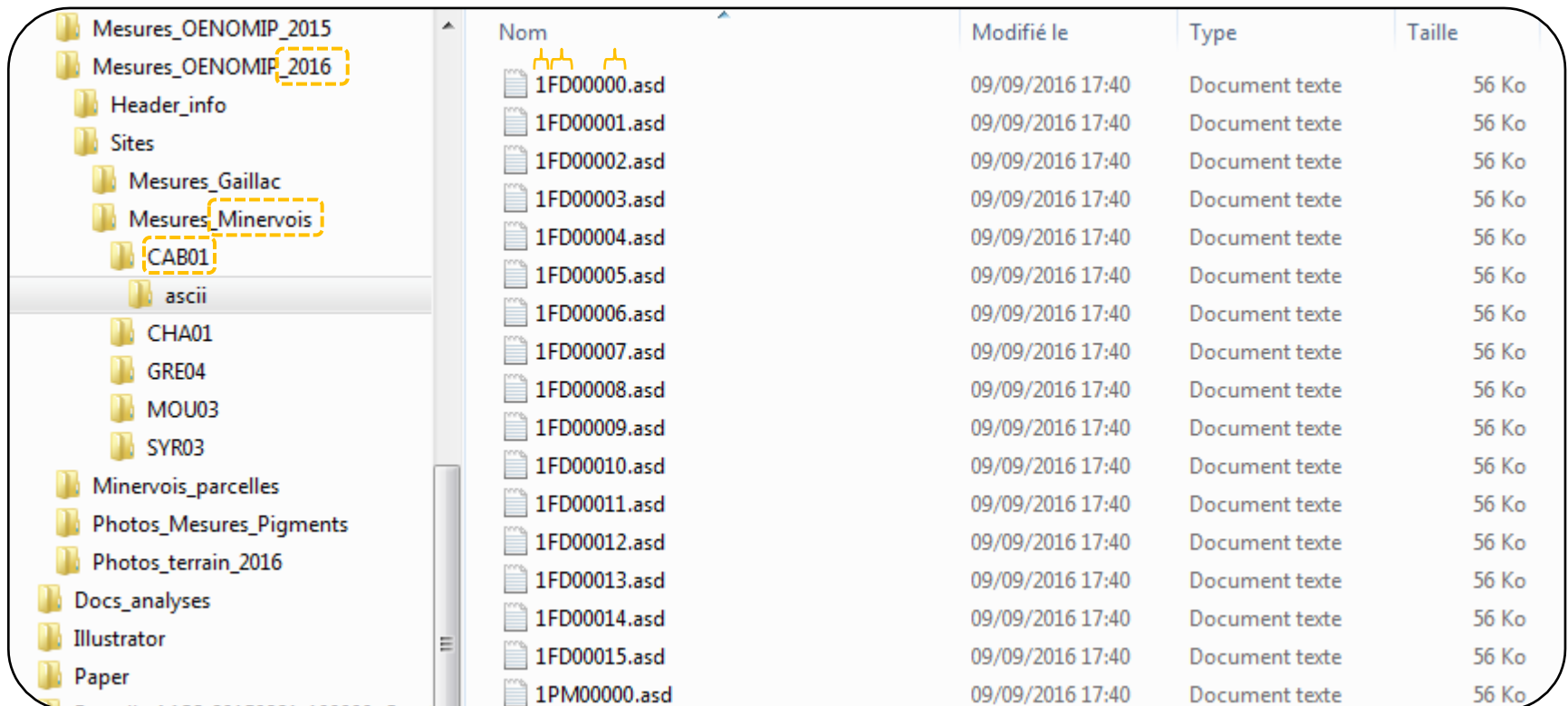
4 types d'erreurs :

- a) Spectres possédant des valeurs  $> 1$
- b) spectres aux valeurs nulles (proche de 0 sur toutes les longueurs d'onde)
- c) **spectres présentant des valeurs de réflectance exceptionnellement faibles** (voire négatives) dans l'intervalle 400-500 nm
- d) spectres présentant une forme à priori normale mais avec des valeurs de réflectance étrangement faibles



# PRÉPARATION BASE DE DONNÉES

Extraire et organiser informations contenues dans les fichiers et dossiers



The image shows a Windows File Explorer window. On the left, the 'Navigation pane' displays a tree of folders. The 'Content pane' on the right shows a list of files with columns for 'Nom', 'Modifié le', 'Type', and 'Taille'. The 'Mesures\_OENOMIP\_2016' folder is selected in the left pane, and its contents are listed in the right pane. The 'ascii' folder is also visible in the left pane.

Nom	Modifié le	Type	Taille
1FD00000.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00001.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00002.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00003.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00004.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00005.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00006.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00007.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00008.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00009.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00010.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00011.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00012.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00013.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00014.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00015.asd	09/09/2016 17:40	Document texte	56 Ko
1PM00000.asd	09/09/2016 17:40	Document texte	56 Ko

# PRÉPARATION BASE DE DONNÉES

Extraire et organiser informations contenues dans les fichiers et dossiers

Nom	Modifié le	Type	Taille
1FD00000.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00001.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00002.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00003.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00004.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00005.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00006.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00007.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00008.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00009.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00010.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00011.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00012.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00013.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00014.asd	09/09/2016 17:40	Document texte	56 Ko
1FD00015.asd	09/09/2016 17:40	Document texte	56 Ko
1PM00000.asd	09/09/2016 17:40	Document texte	56 Ko

**Attributs extraits directement :**

- Année, Site, Id parcelle, Id pied (terrain), Etat de santé, Numéro de la mesure

**Indirectement :**

- Nom du cépage, couleur du cépage, Id feuille

# PRÉPARATION BASE DE DONNÉES

Extraire et organiser informations contenues dans les fichiers et dossiers

Aperçu de la base de données (stockée en .HDF5 => taille sur disque ~100 Mo)

## Matrice X : Spectres

Index	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
0	0.0569	0.0565	0.0513	0.0475	0.0457	0.0454	0.0445	0.044	0.0461	0.0424	0.0399	0.041	0.0406	0.0413	0.0427	0.0435	0.0426	0.0413	0.0401	0.0417	0.0418	0.0401	0.0408	0.0416	0.0421	0.0417
1	0.0637	0.0632	0.0568	0.0519	0.0493	0.0496	0.0494	0.0489	0.0484	0.0452	0.0439	0.0464	0.045	0.0443	0.0443	0.045	0.0438	0.0423	0.0421	0.0424	0.0424	0.042	0.0422	0.0438	0.0452	0.0429
2	0.0705	0.0703	0.0639	0.0589	0.056	0.0568	0.0576	0.0576	0.0547	0.052	0.0518	0.0552	0.0546	0.0534	0.0522	0.0519	0.0512	0.0504	0.0495	0.0494	0.0496	0.0498	0.0488	0.0491	0.0499	0.0499
3	0.0668	0.0665	0.0596	0.0538	0.05	0.05	0.0489	0.0477	0.0483	0.045	0.0439	0.0473	0.0465	0.0455	0.0449	0.0463	0.0456	0.0443	0.0437	0.0437	0.0437	0.0438	0.0438	0.0437	0.0437	0.0439
4	0.0596	0.0586	0.0512	0.0478	0.0487	0.0485	0.0476	0.047	0.0486	0.0447	0.0426	0.0451	0.0453	0.0444	0.0431	0.0434	0.0419	0.0405	0.0409	0.0413	0.0414	0.0412	0.0411	0.0412	0.0415	0.042
5	0.0671	0.0667	0.0596	0.0559	0.0562	0.0542	0.053	0.0526	0.0533	0.0498	0.048	0.0503	0.0509	0.05	0.0483	0.0481	0.0469	0.046	0.0468	0.0468	0.0466	0.0462	0.0454	0.0457	0.0465	0.0466

## Matrice Y : Attributs

Index	parcel_code	health_state	plant_number	measure_number	cepage_code	cepage_name	vine_color	vine_color_code	year	site	plant_id_str	plant_id_int	leaf_id
0	1	2	1	10	1	Rouge Braucol	Rouge	1	2015	Gaillac	1_2_1	0	3
1	1	2	6	12	1	Rouge Braucol	Rouge	1	2015	Gaillac	1_2_6	1	4
2	1	2	5	1	1	Rouge Braucol	Rouge	1	2015	Gaillac	1_2_5	2	1
3	1	1	4	3	1	Rouge Braucol	Rouge	1	2015	Gaillac	1_1_4	3	1
4	1	2	4	4	1	Rouge Braucol	Rouge	1	2015	Gaillac	1_2_4	4	2
5	1	1	6	13	1	Rouge Braucol	Rouge	1	2015	Gaillac	1_1_6	5	4

# PRÉPARATION BASE DE DONNÉES

## Interroger la base de données

Calculer statistiques descriptives  
et générer des tableaux

Index	Num_plants	Num_mesures
Blanc Cabernet Sauvignon	21	337
Blanc Chardonnay	18	288
Blanc Colombard	20	320
Blanc Loin de l Oeil	20	320
Blanc Muscadelle	40	641
Blanc Ondenc	30	475
Rouge Braucol	50	800
Rouge Cabernet Sauvignon	30	464
Rouge Duras	50	798
Rouge Gamay	40	638
Rouge Grenache	30	480
Rouge Merlot	20	311
Rouge Mourvedre	30	480
Rouge Syrah	24	375



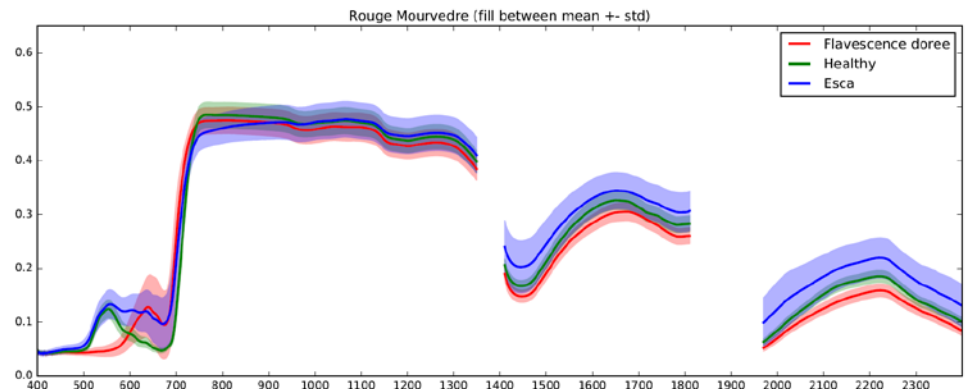
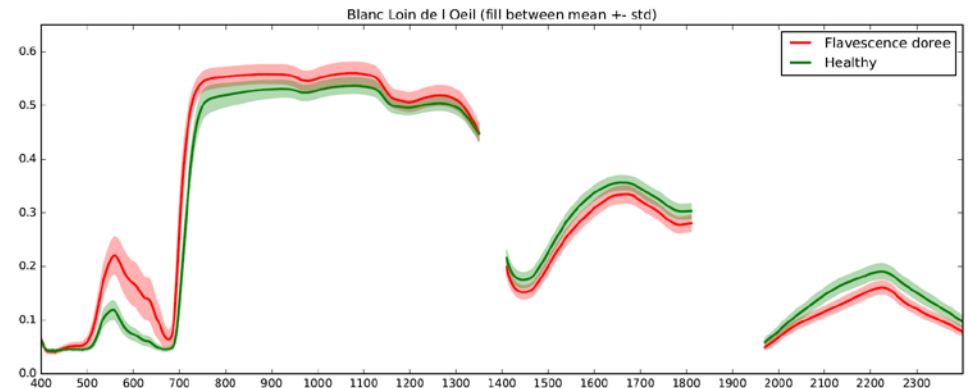
# PRÉPARATION BASE DE DONNÉES

## Interroger la base de données

Calculer statistiques descriptives  
et générer des tableaux

Index	Num_plants	Num_mesures
Blanc Cabernet Sauvignon	21	337
Blanc Chardonnay	18	288
Blanc Colombard	20	320
Blanc Loin de l Oeil	20	320
Blanc Muscadelle	40	641
Blanc Ondenc	30	475
Rouge Braucol	50	800
Rouge Cabernet Sauvignon	30	464
Rouge Duras	50	798
Rouge Gamay	40	638
Rouge Grenache	30	480
Rouge Merlot	20	311
Rouge Mourvedre	30	480
Rouge Syrah	24	375

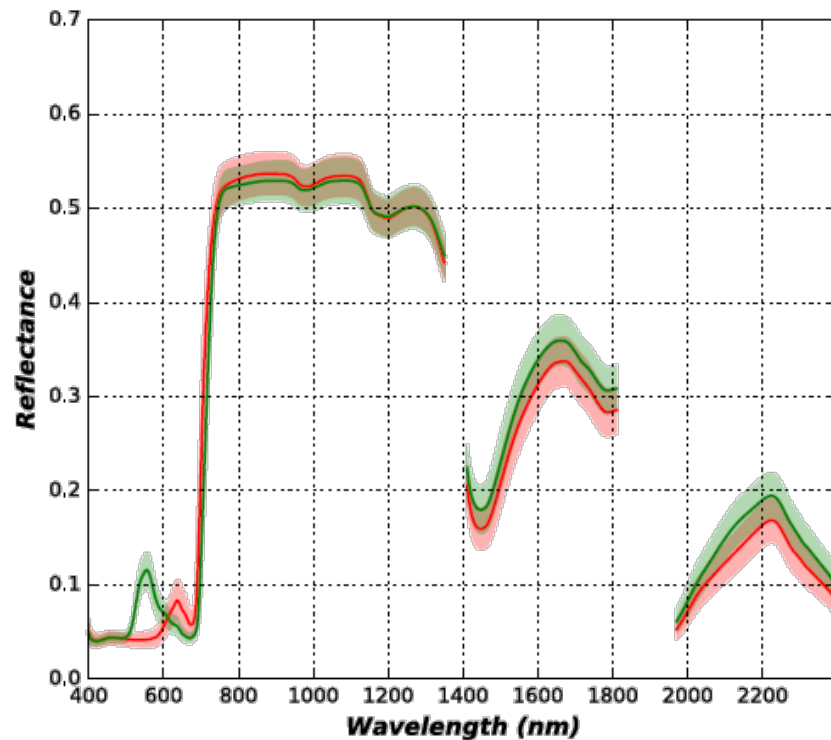
Tracer courbes spectrales  
(moyennes, std, min, max, etc.)



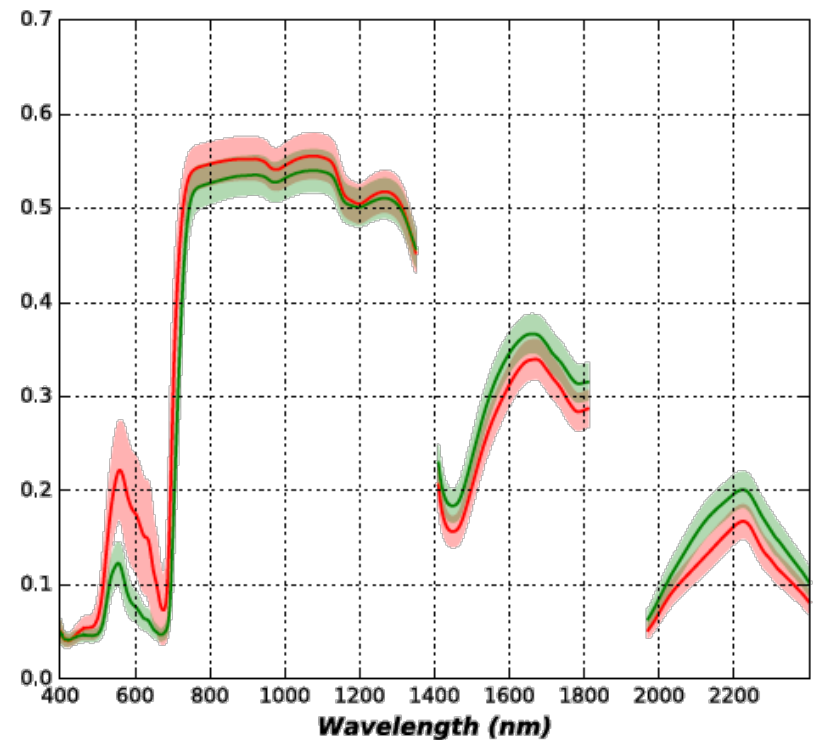
# PRÉPARATION BASE DE DONNÉES

Jeu de données 2015

a) Red dataset

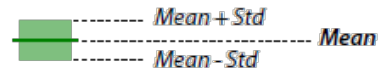


b) White dataset

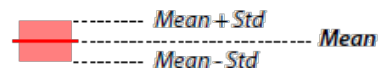


Spectral  
signatures

Healthy



Flavescente dorée



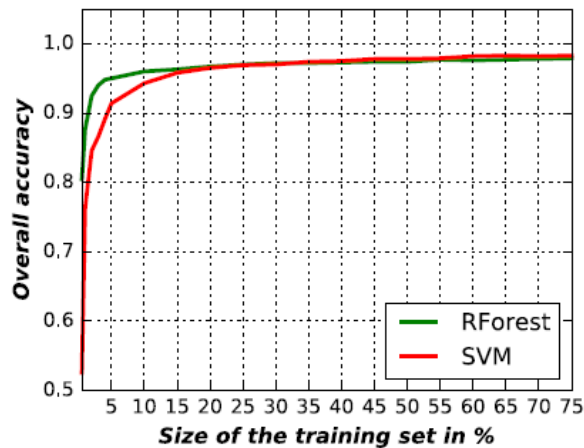
# EXPÉS - CLASSIFICATIONS AVEC LA TOTALITÉ DE L'INFO SPECTRALE

Données 2015 - 2 classes FD et PS

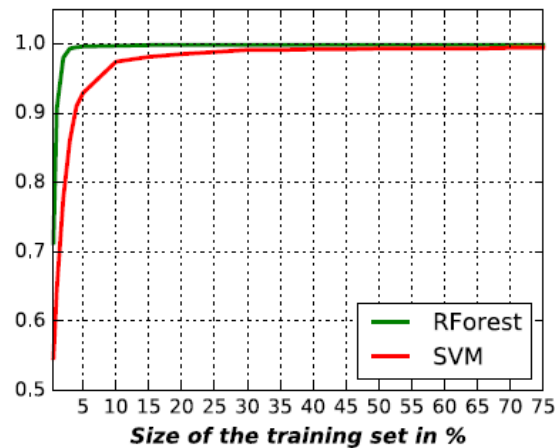
*Baselines avec Random Forest et SVM*

- Qualité des prédictions en fonction de la taille du *training set*

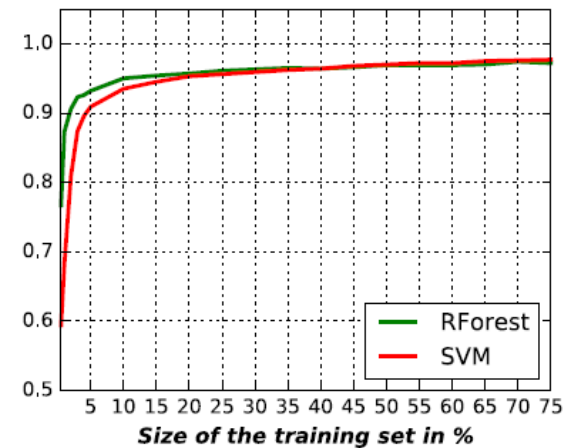
Red & White



Red



White



# EXPÉS - CLASSIFICATIONS AVEC LA TOTALITÉ DE L'INFO SPECTRALE

**Données 2015** - 2 classes FD et PS

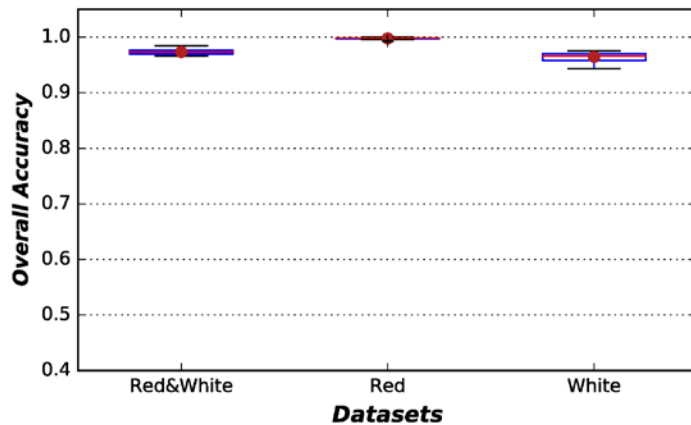
**Evaluation de l'impact de l'échantillonnage sur la qualité des prédictions**  
(résultats affichés uniquement *Random Forest*)

4 types d'échantillonnage pour construire : *training set* / *test set*

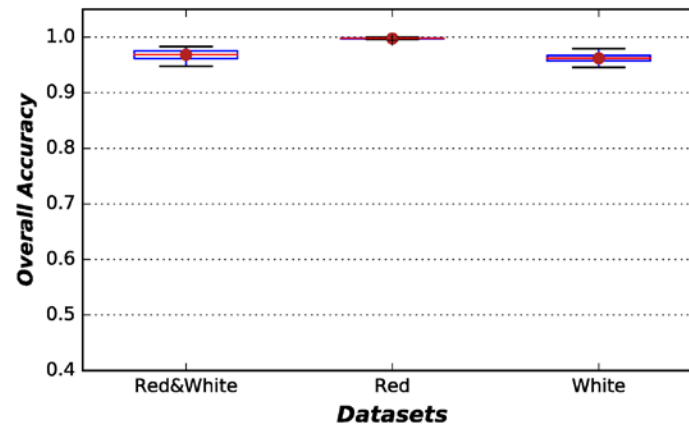
- a) pas de regroupement des spectres
- b) spectres regroupés par feuille
- c) spectres regroupés par pied de vigne
- d) spectres regroupés par cépage

# EXPÉS - CLASSIFICATIONS AVEC LA TOTALITÉ DE L'INFO SPECTRALE

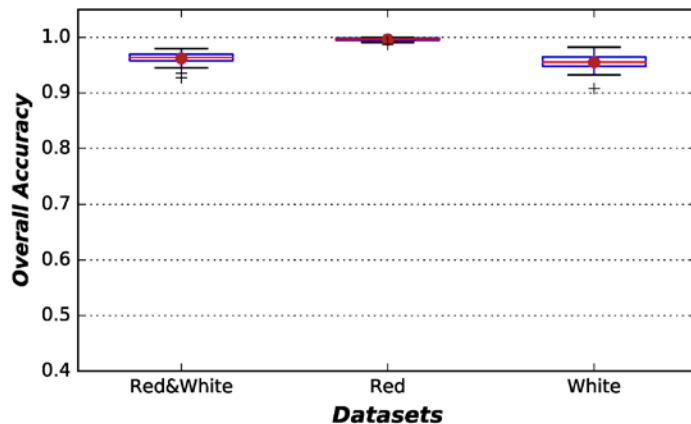
a) *NotGrouped*



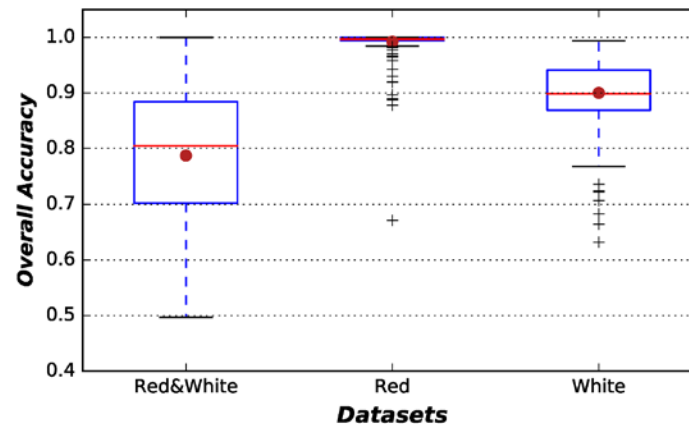
b) *GroupedByLeaf*



c) *GroupedByVine*



d) *GroupedByCultivar*



# EXPÉS - SÉLECTION DES PORTIONS SPECTRALES LES PLUS PERTINENTES

**Données 2015**

**Sélection automatisée d'un nombre réduit de bandes** permettant d'atteindre une qualité de prédiction comparable à la *baseline* (*R. Forest full spectral range*)

2 stratégies de sélection de bandes :

a) *Ranking of Importance* (RI) couplé au classifieur *Random Forest* (RF)

**RI-RF**

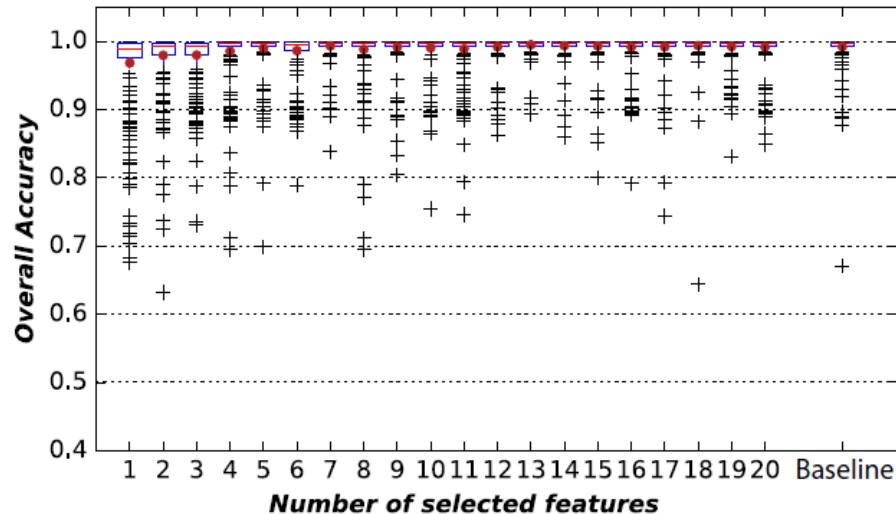
b) *Sequential Forward Selection* (SFS) couplé au classifieur *Gaussian Mixture Models* (GMM)

**SFS-GMM**

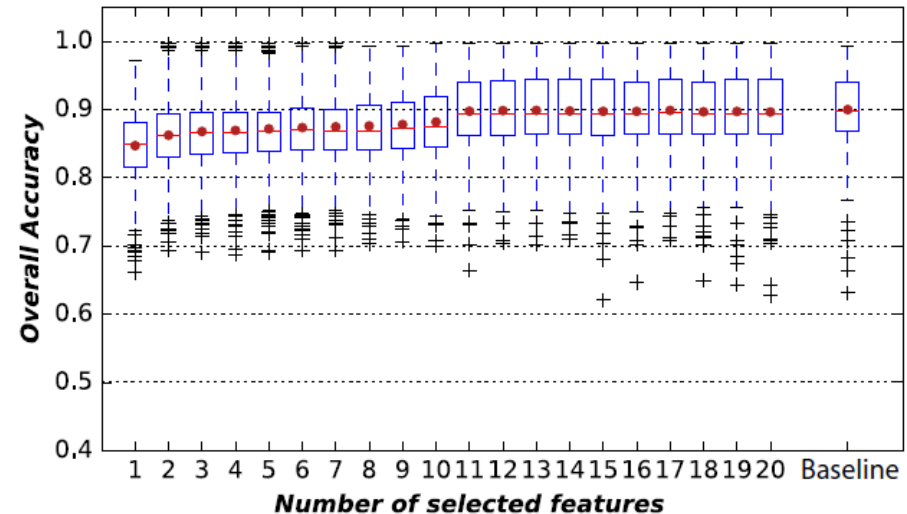
# EXPÉS - SÉLECTION DES PORTIONS SPECTRALES LES PLUS PERTINENTES

Détermination du nombre optimal de bandes via RI-RF

a) *Red* dataset with RI-RF



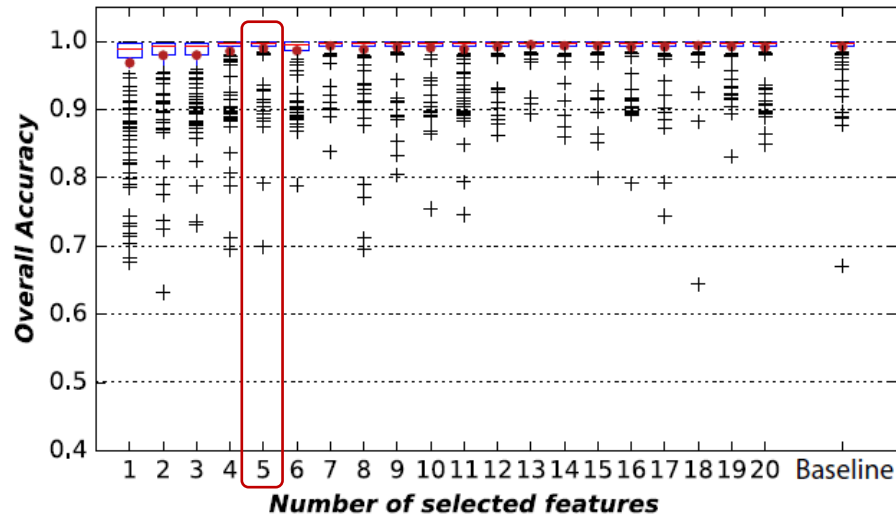
b) *White* dataset with RI-RF



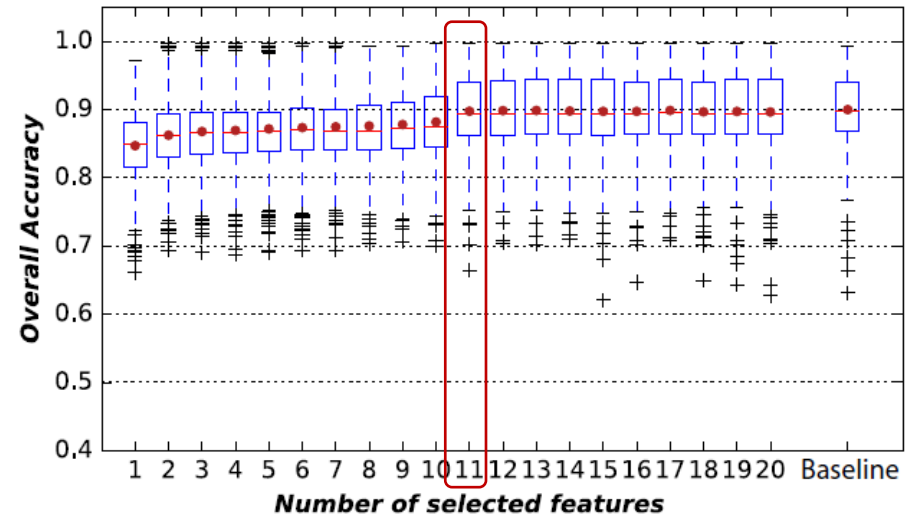
# EXPÉS - SÉLECTION DES PORTIONS SPECTRALES LES PLUS PERTINENTES

Détermination du nombre optimal de bandes via RI-RF

a) *Red* dataset with RI-RF



b) *White* dataset with RI-RF

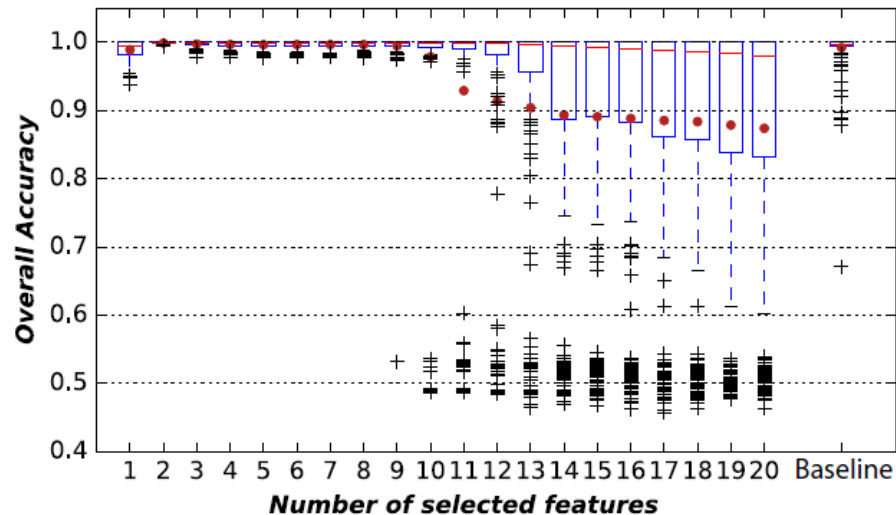




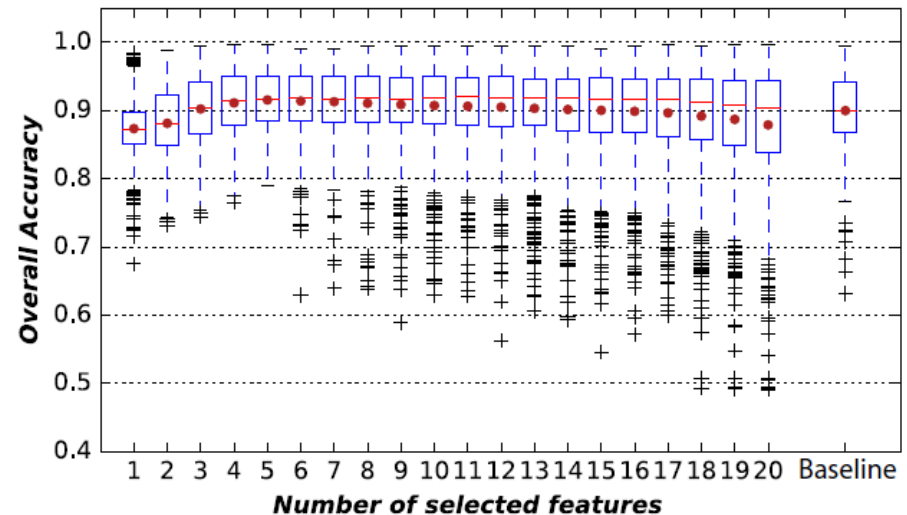
# EXPÉS - SÉLECTION DES PORTIONS SPECTRALES LES PLUS PERTINENTES

Détermination du nombre optimal de bandes via SFS-GMM

c) *Red* dataset with SFS-GMM



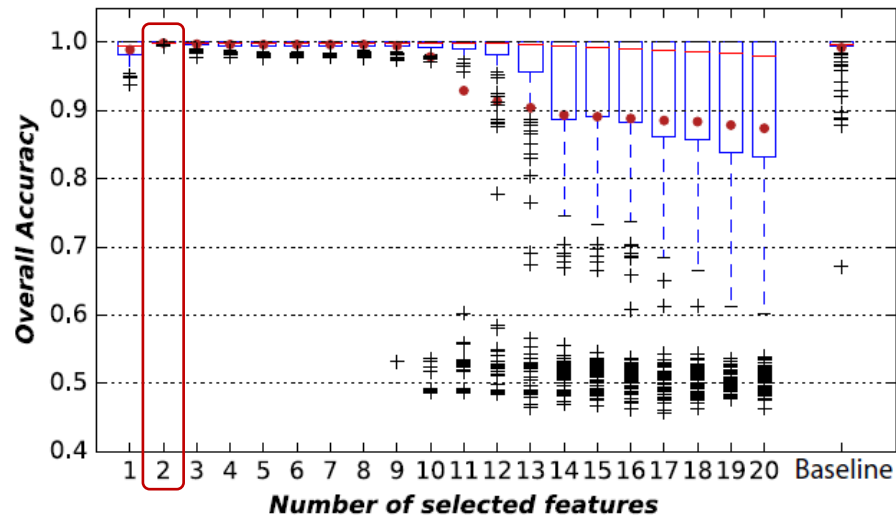
d) *White* dataset with SFS-GMM



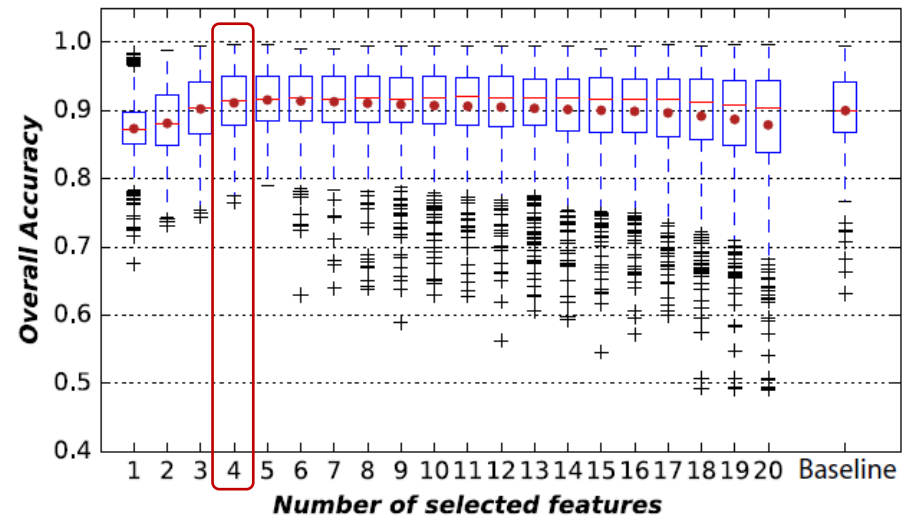
# EXPÉS - SÉLECTION DES PORTIONS SPECTRALES LES PLUS PERTINENTES

Détermination du nombre optimal de bandes via SFS-GMM

c) *Red* dataset with SFS-GMM



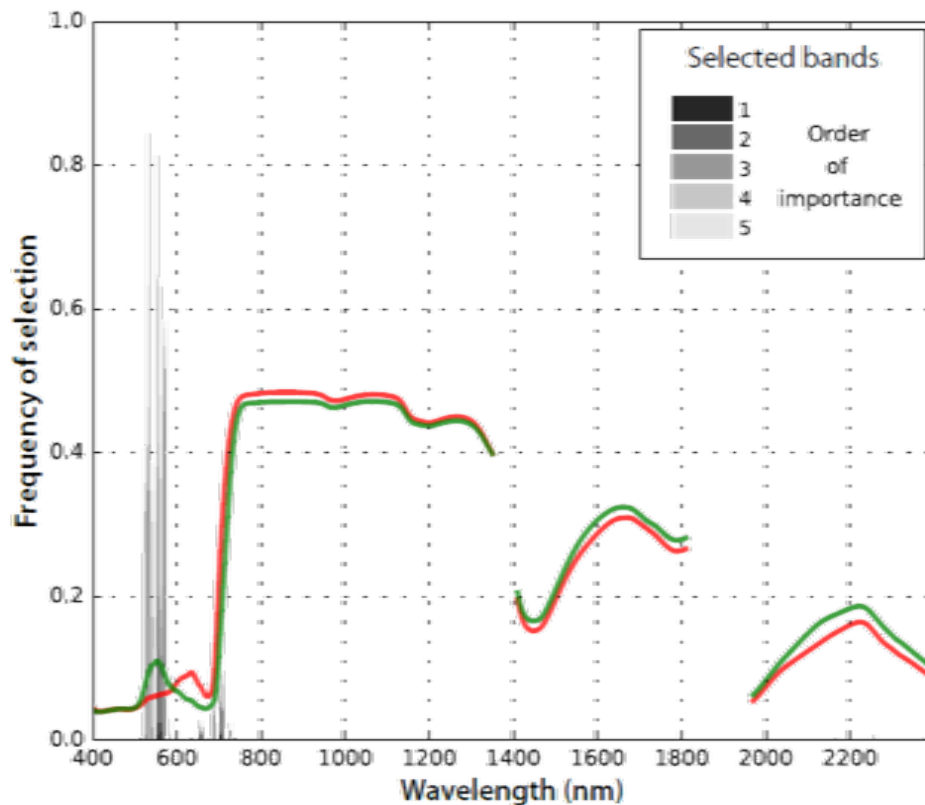
d) *White* dataset with SFS-GMM



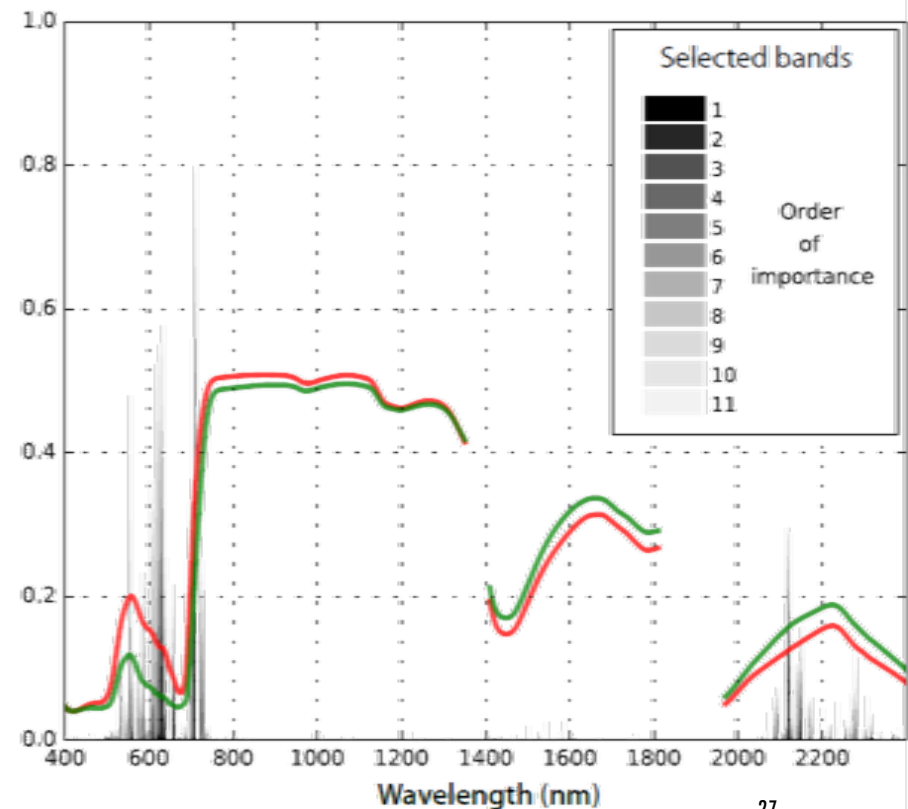
# EXPÉS - SÉLECTION DES PORTIONS SPECTRALES LES PLUS PERTINENTES

Bandes sélectionnées via RI-RF

a) *Red* dataset with RI-RF



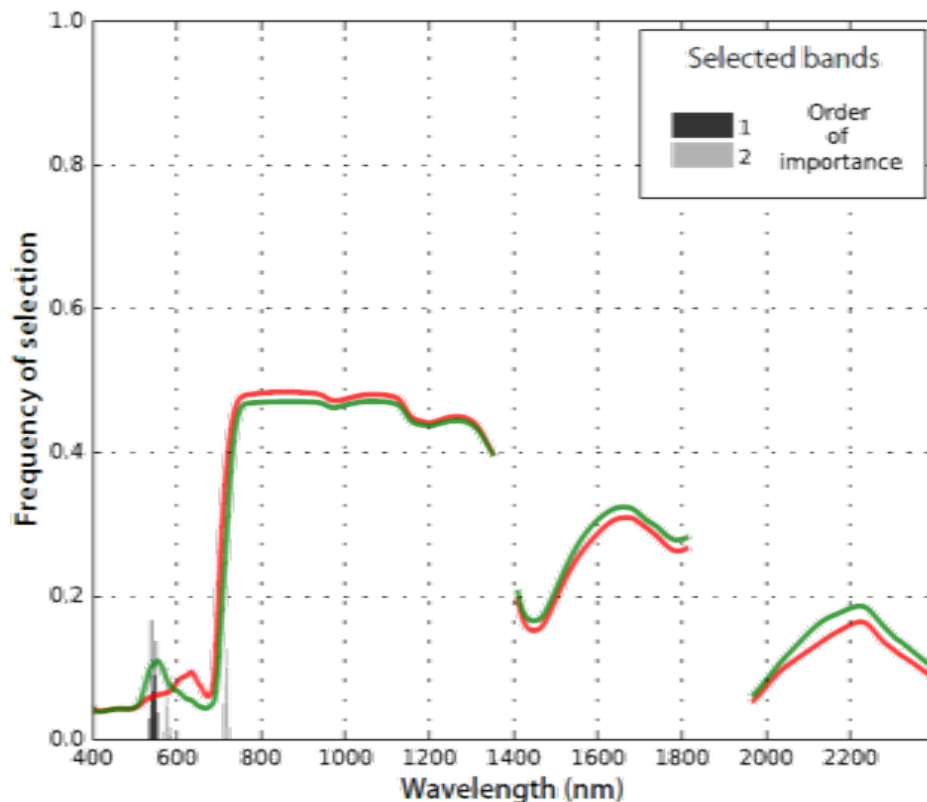
b) *White* dataset with RI-RF



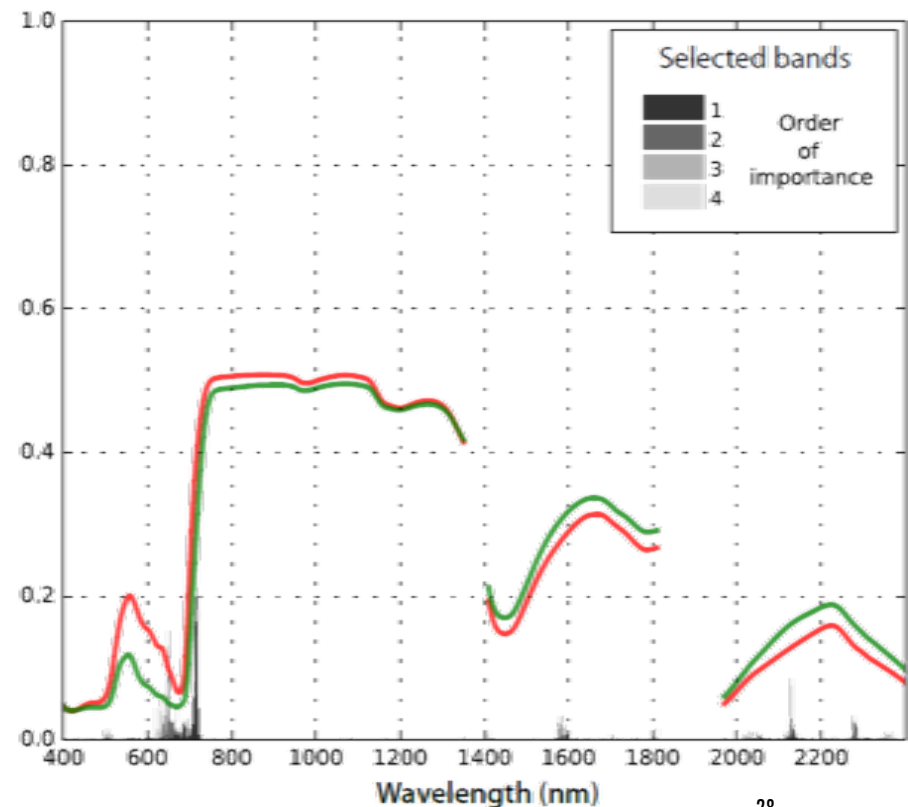
# EXPÉS - SÉLECTION DES PORTIONS SPECTRALES LES PLUS PERTINENTES

Bandes sélectionnées via SFS-GMM

c) Red dataset with SFS-GMM



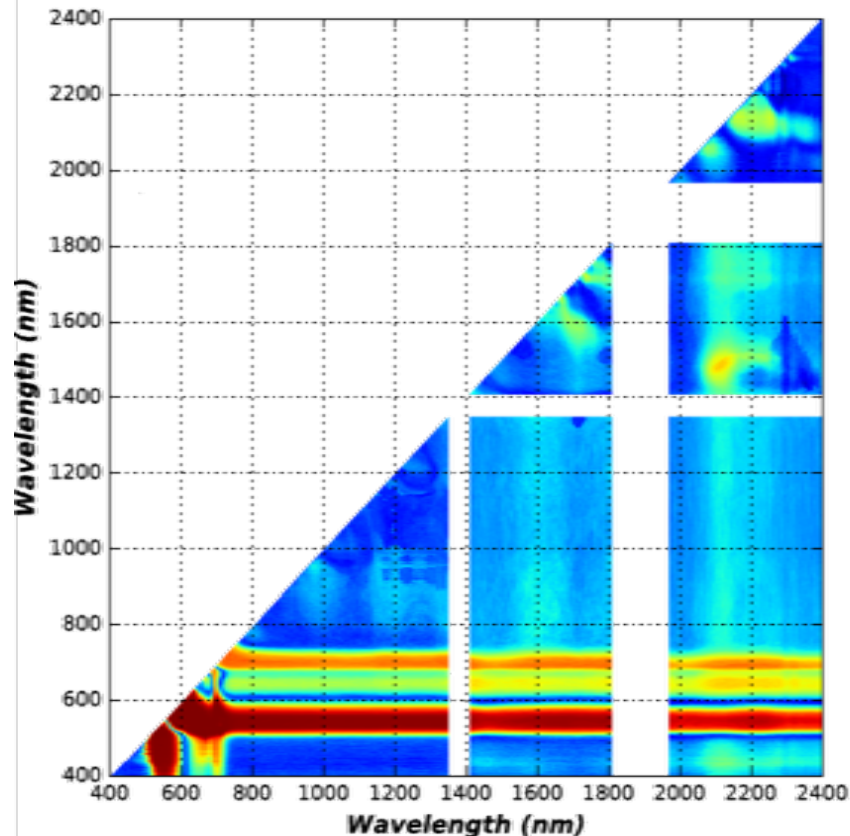
d) White dataset with SFS-GMM



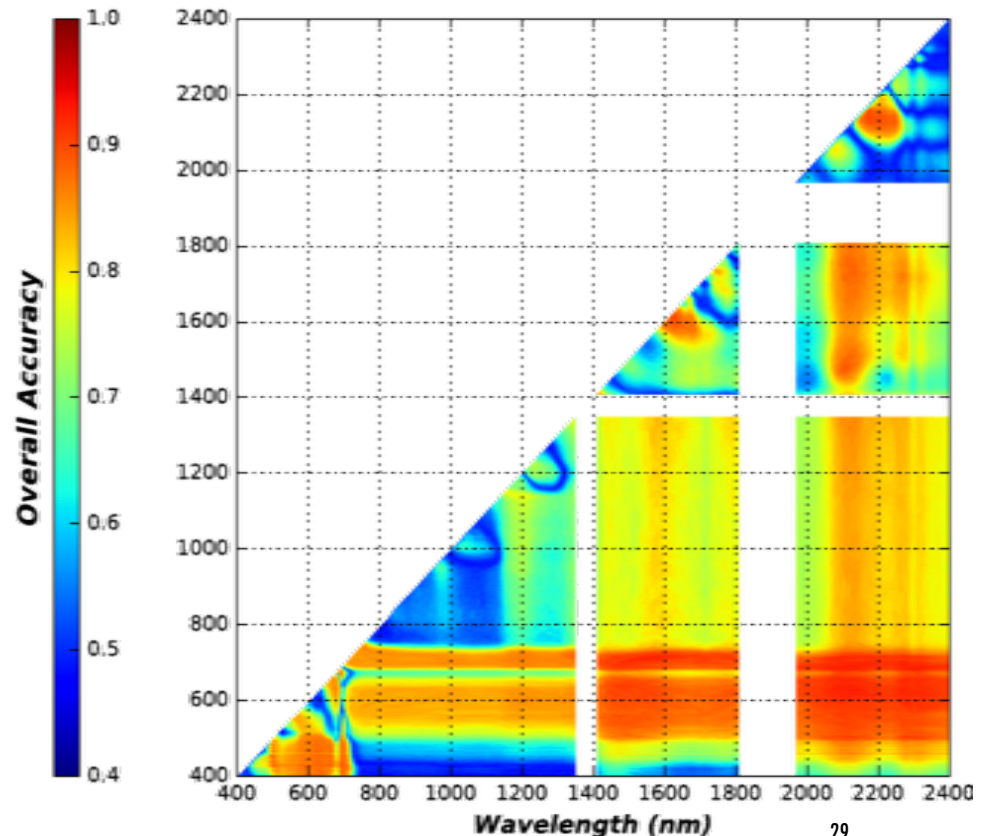
# EXPÉS - SÉLECTION DES PORTIONS SPECTRALES LES PLUS PERTINENTES

Sélection d'un indice normalisé optimal

a) Red dataset with RF



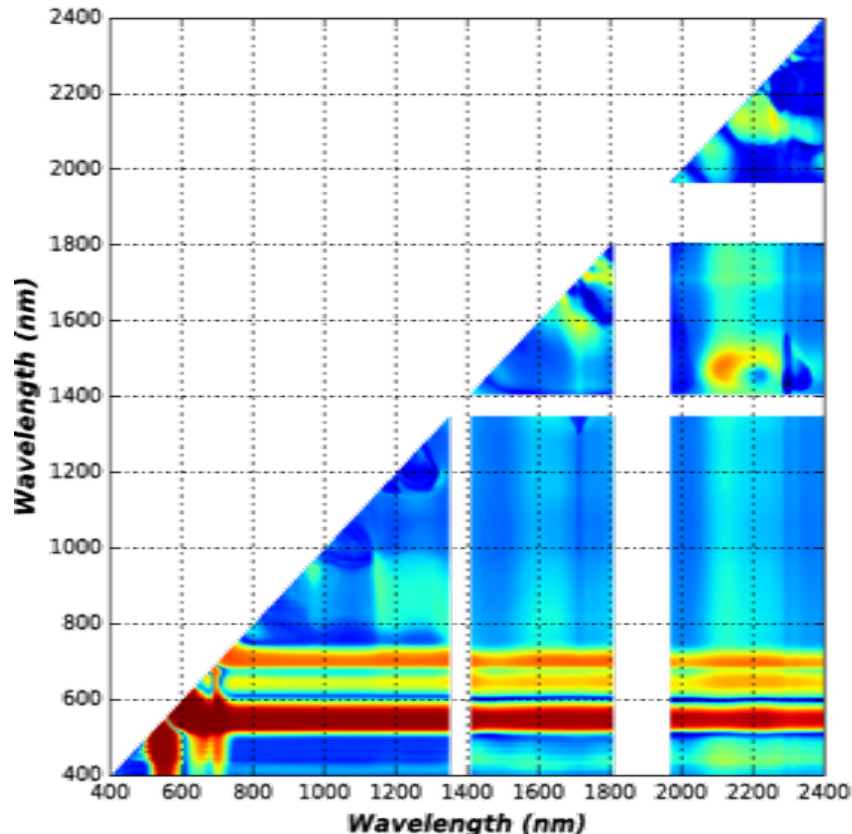
b) White dataset with RF



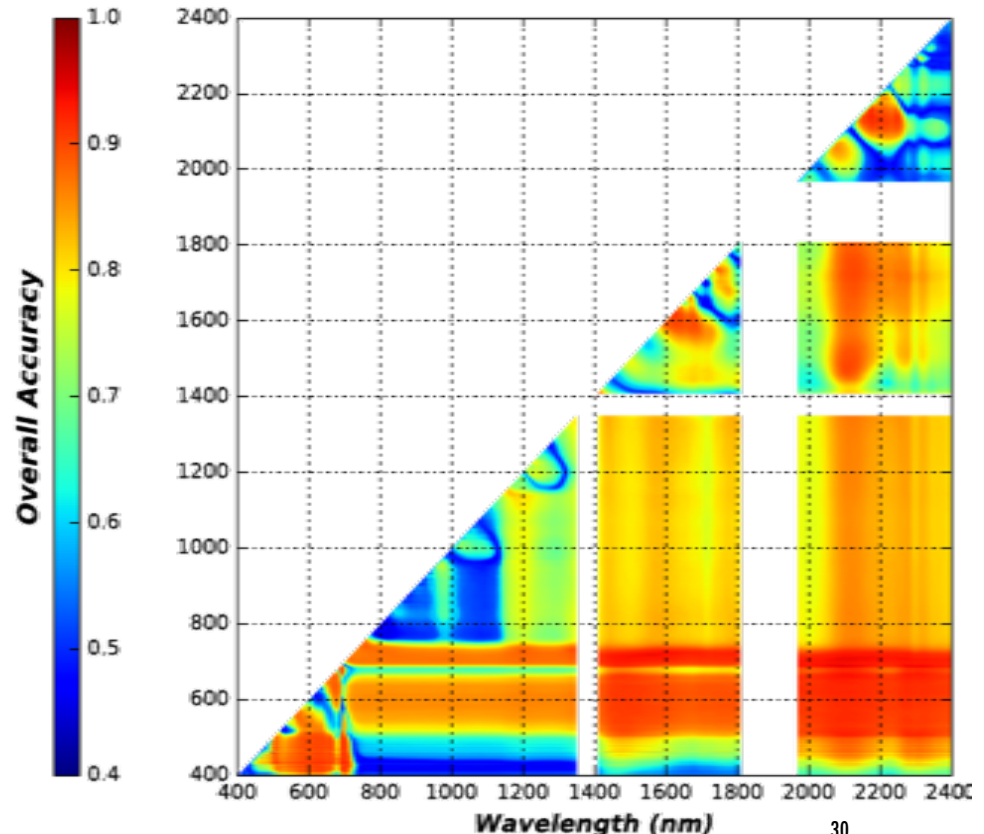
# EXPÉS - SÉLECTION DES PORTIONS SPECTRALES LES PLUS PERTINENTES

Sélection d'un indice normalisé optimal

c) *Red* dataset with GMM



d) *White* dataset with GMM



# EXPÉS - SIMULATION DE CAPTEURS (DRONE) MULTISPECTRAUX

**Simuler le signal enregistré par des capteurs multispectraux à partir de la collection de spectres**

**Prise en compte uniquement de la sensibilité spectrale et de la résolution radiométrique des capteurs**

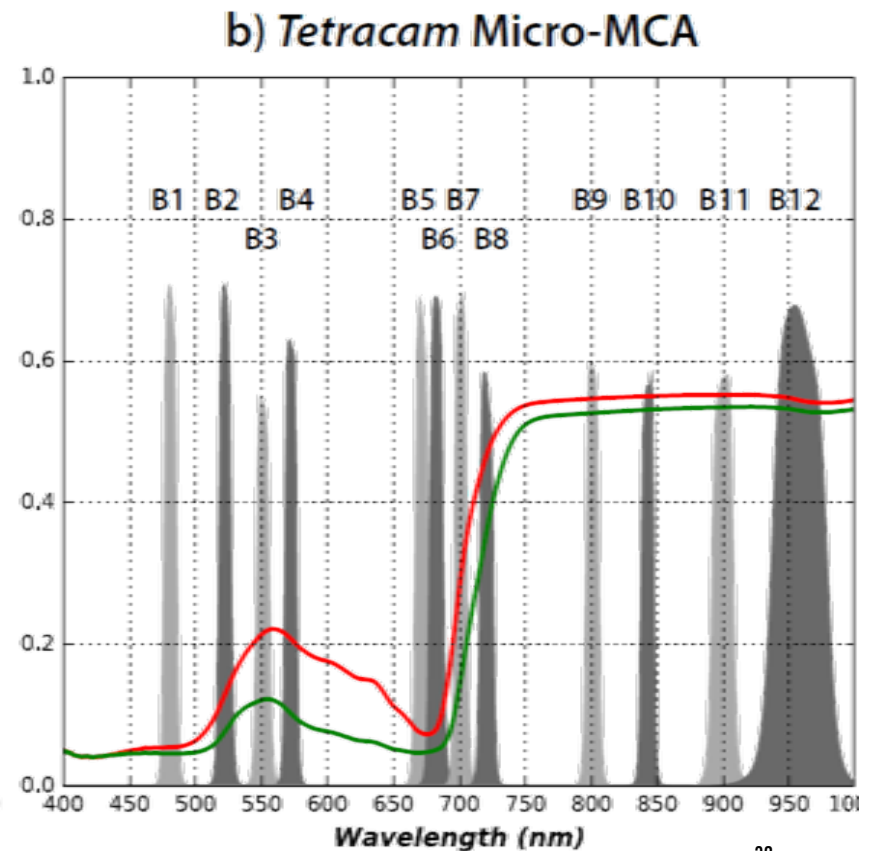
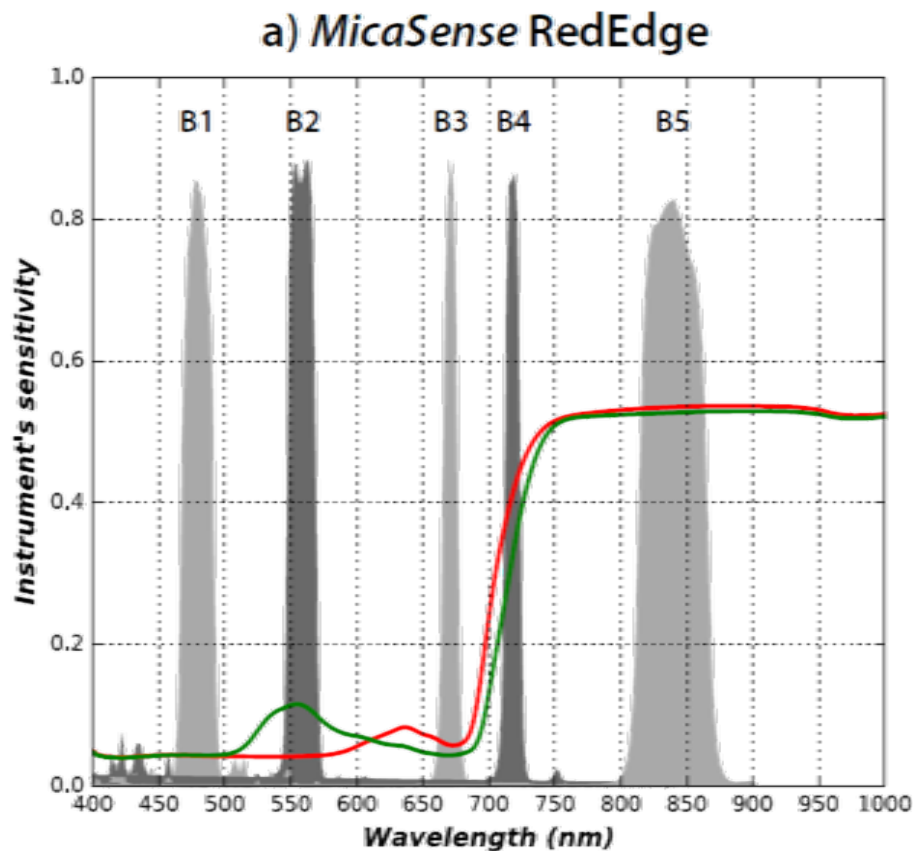
**Capteurs simulés :**

**a) *MicaSense RedEdge* 5 bandes (fixes)**

**b) *Tetracam Micro-MCA* 4, 6 ou 12 bandes (ajustables)**

# EXPÉS - SIMULATION DE CAPTEURS (DRONE) MULTISPECTRAUX

Capteurs simulés





# EXPÉS - SIMULATION DE CAPTEURS (DRONE) MULTISPECTRAUX

Réalisation de la même chaîne d'expés sur données simulés

a) *MicaSense RedEdge* 5 bandes

b) *Tetracam Micro-MCA* 12 bandes

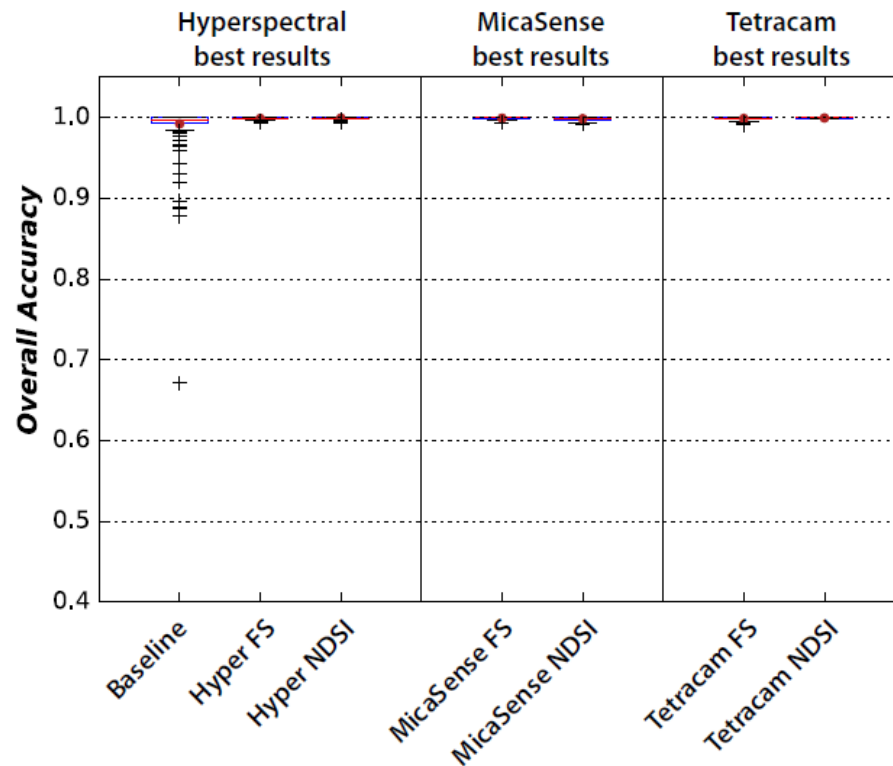
Sélection nombre optimal de bandes

Fréquence de sélection des bandes

Evaluation exhaustive des indices normalisés

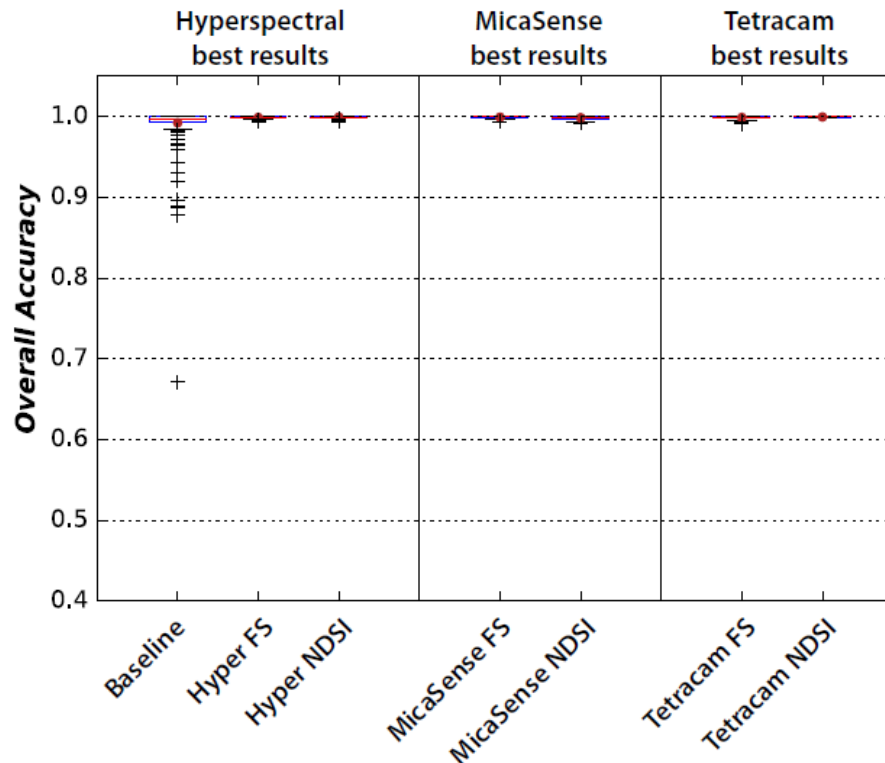
# EXPÉS - COMPARAISONS - CONCLUSIONS

a) *Red* dataset



# EXPÉS - COMPARAISONS - CONCLUSIONS

a) *Red* dataset



**Hyper FS** - 2 bandes parmi :  
~ 545, 710, 575 nm

**Mica FS** - 2 bandes :  
Green et RedEdge

**Tetra FS** - 2 bandes parmi :  
B3 (550), B8 (720) et B4 (570)

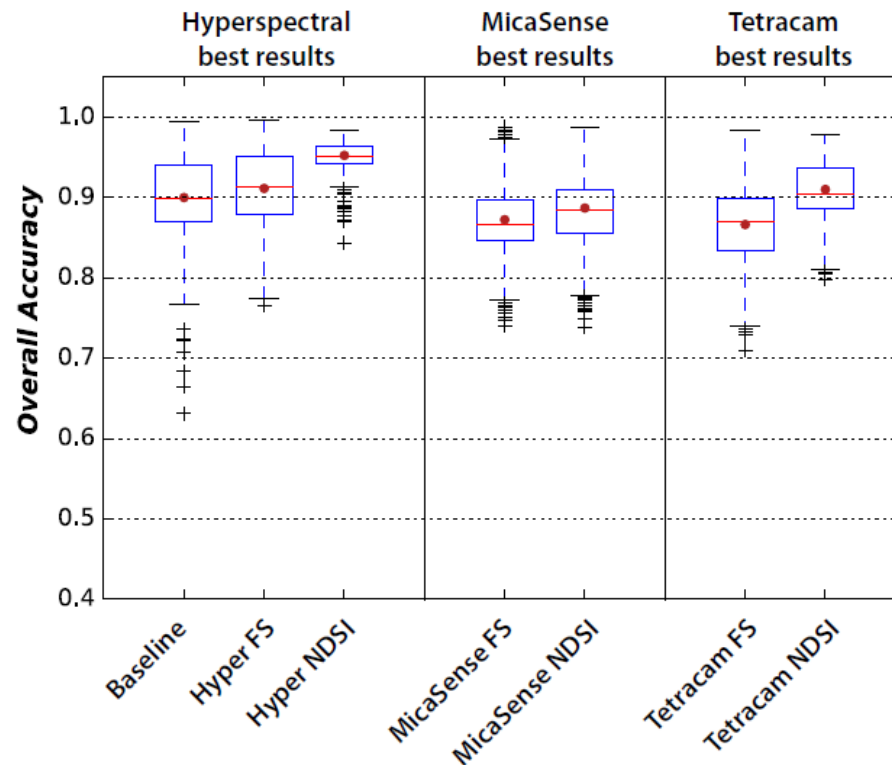
**Hyper NDSI** : 550 et 610 nm

**Mica NDSI** : Green et Red

**Tetra NDSI** : B4 (570) et B7 (700)

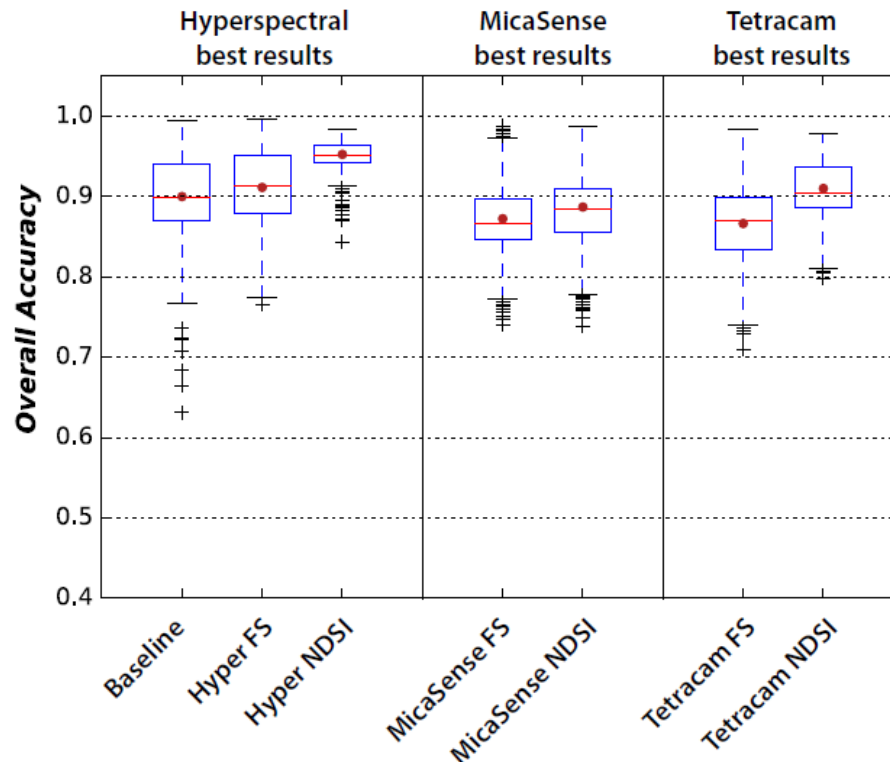
# EXPÉS - COMPARAISONS - CONCLUSIONS

## b) *White* dataset



# EXPÉS - COMPARAISONS - CONCLUSIONS

## b) *White* dataset



**Hyper FS** - 4 bandes parmi :

~ 710, 690, 2135, 645, 2280 nm

**Hyper NDSI** : 687 et 2105 nm

**Mica FS** - 1 bande parmi :

RedEdge, Green, Red

**Mica NDSI** : RedEdge et NIR

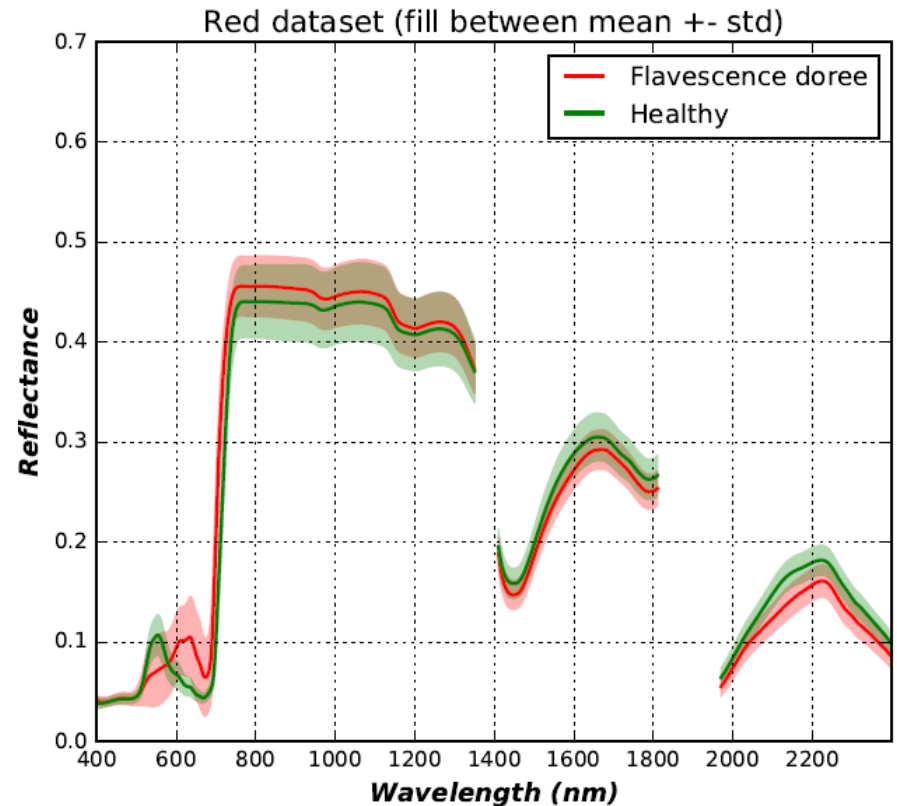
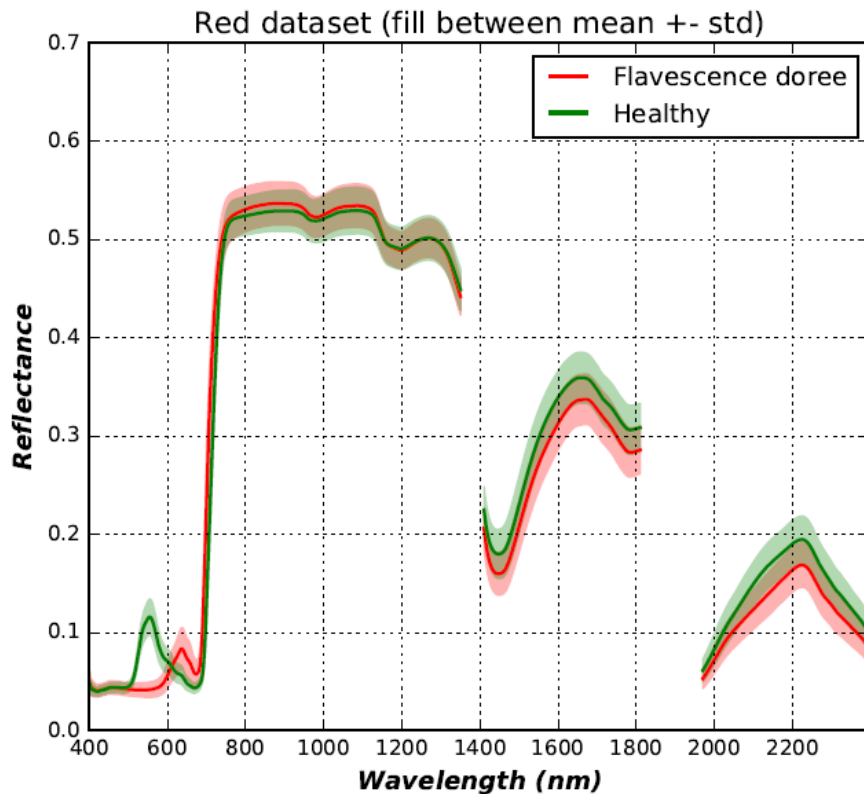
**Tetra FS** - 1 bande parmi :

B8 (720), B7 (700), B4 (570)

**Tetra NDSI** : B1 (480) et B7 (700)

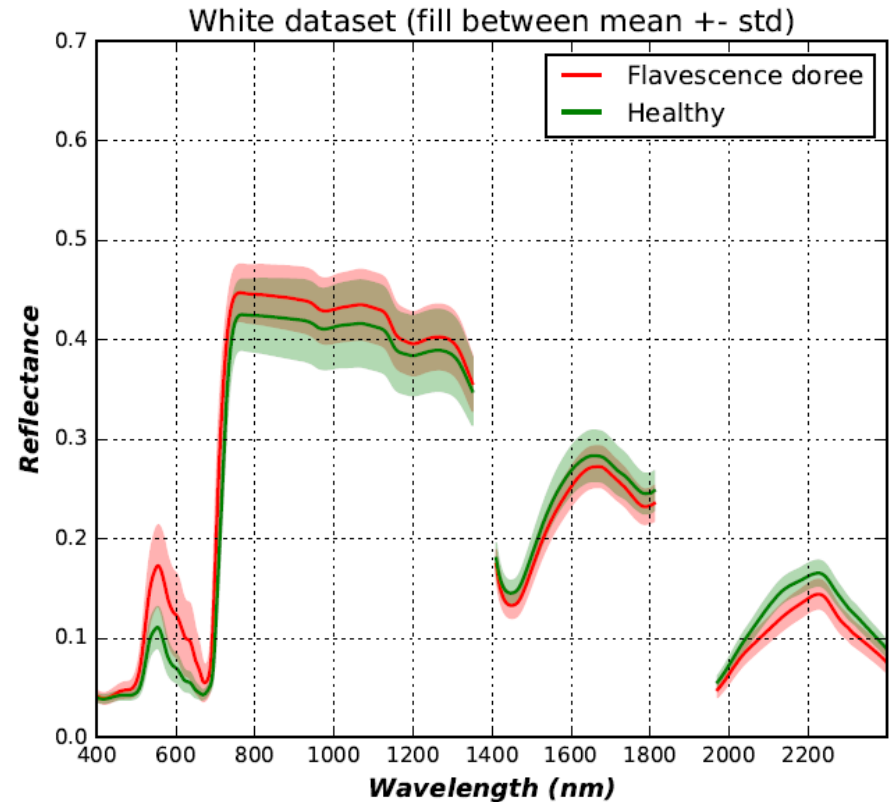
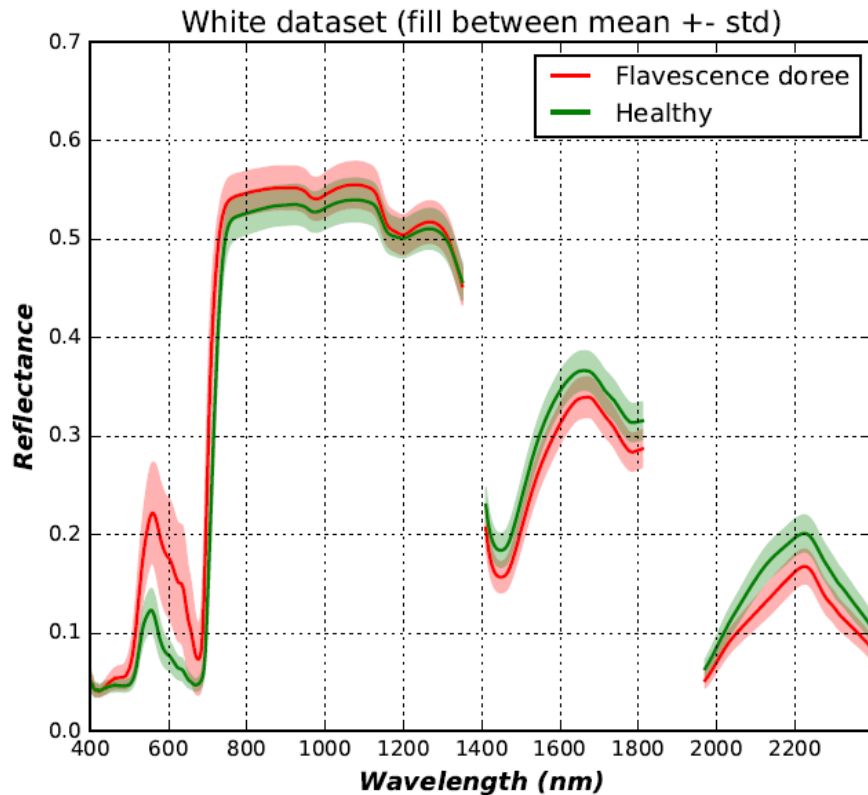
# COMPARAISONS 2015 - 2016

## Courbes spectrales



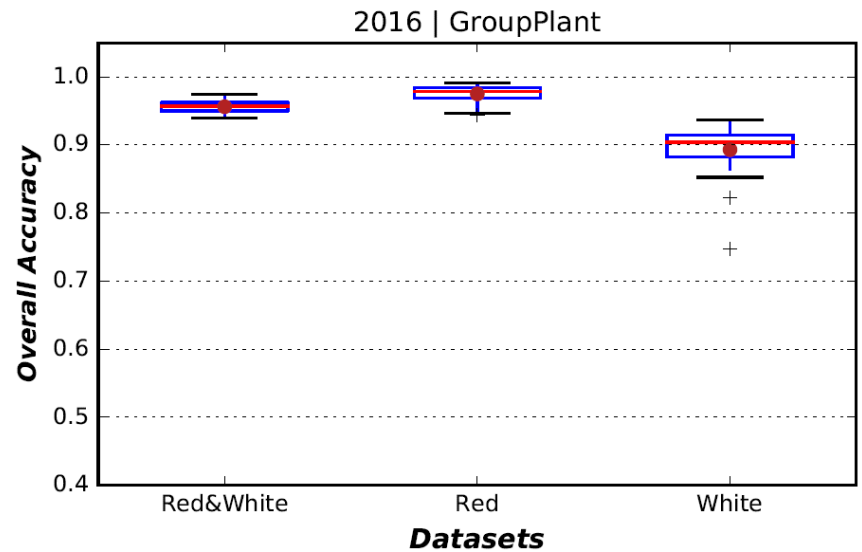
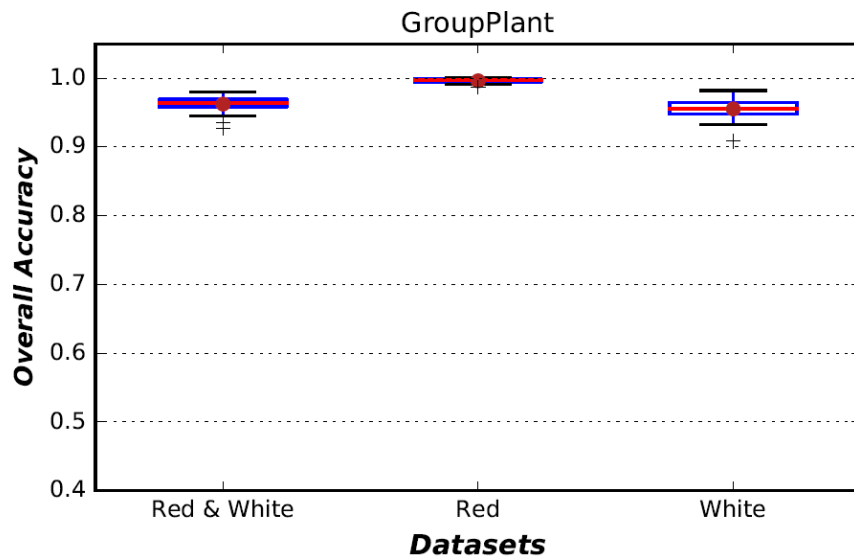
# COMPARAISONS 2015 - 2016

## Courbes spectrales



# COMPARAISONS 2015 - 2016

## Qualité des prédictions





# COMPARAISONS 2015 - 2016

## Qualité des prédictions

