

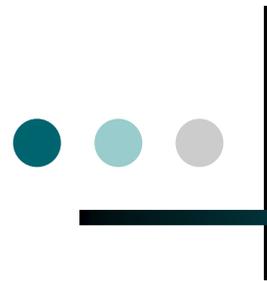
En quoi l'analyse multi-tableaux peut-elle être utile ?



Conférence Héliospir – 28/08/08

S.Preys – Ondalys - Montpellier

G.Mazerolles – UMR SPO – INRA - Montpellier



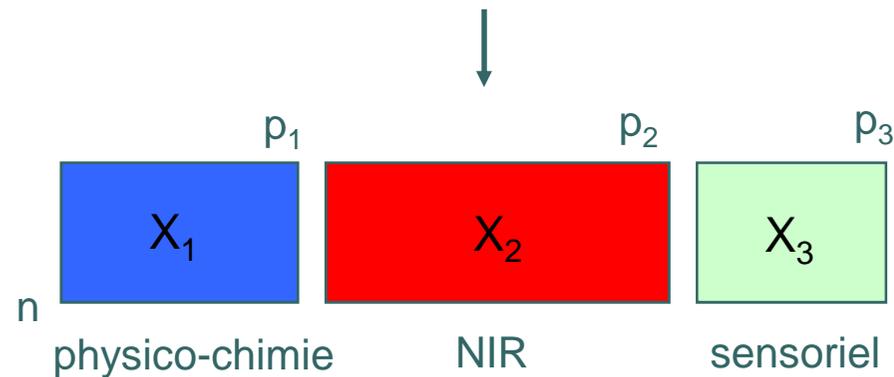
Sommaire

- I. Problématique**
- II. Théorie
- III. Exemple d'application
- IV. Conclusions



Généralités

Différents types de mesures multivariées sur les mêmes individus



Multi-tableaux = tableaux multiples = multi-blocs / *Multiblock* = *multiset* = *K-tables*



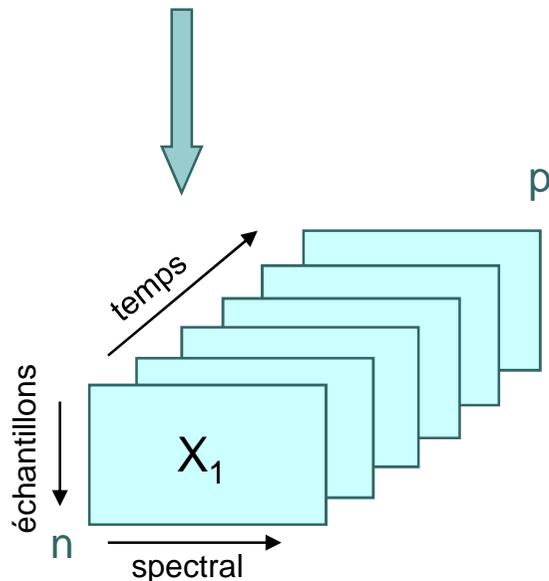
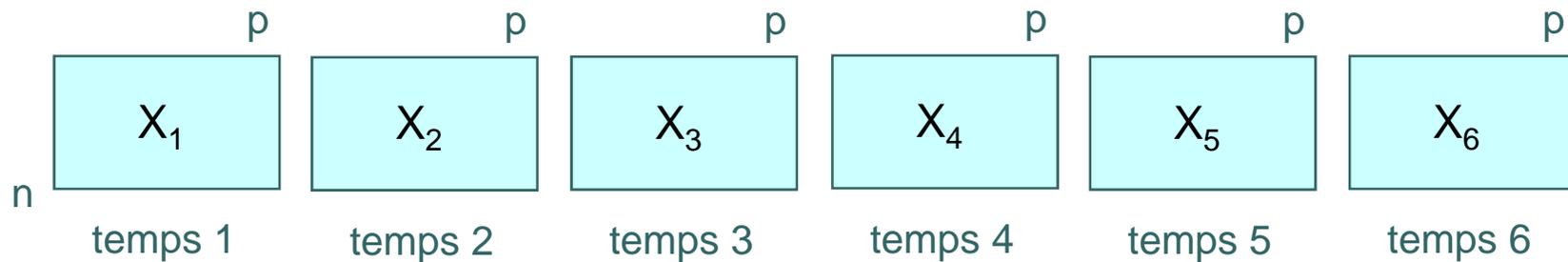
Quelle analyse de données ?

Quels objectifs ?



Types de données

- Mesures de spectres dans le temps :



Multiway / Multi-voies = cubes ou hypercubes

Pas multi-tableaux !

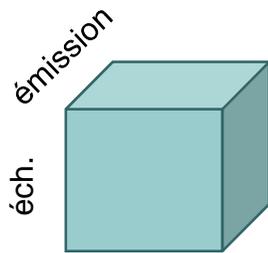
Analyses multiway :

- PARAFAC
- TUCKER

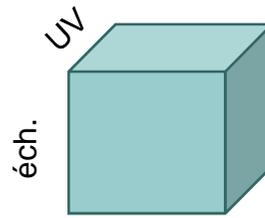


Types de données

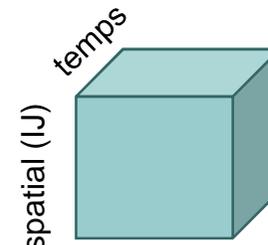
- Autres exemples de données multiway (cubes) :



Fluorescence



Chromatographie

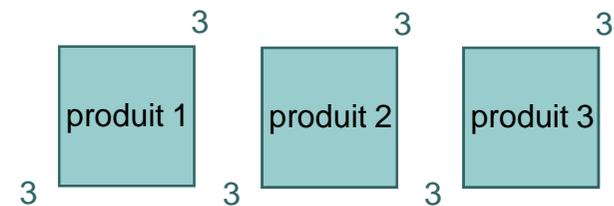
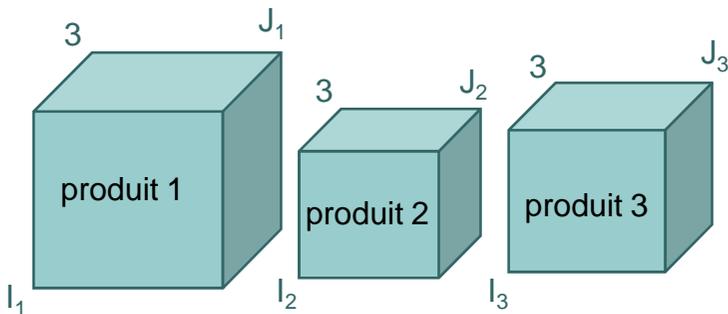
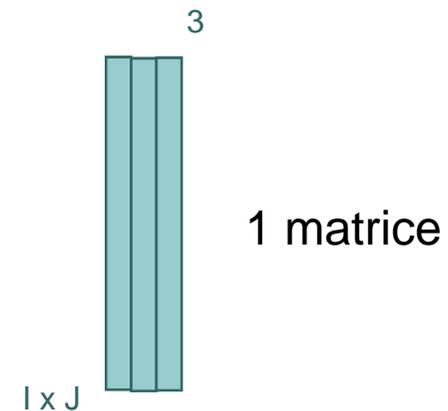
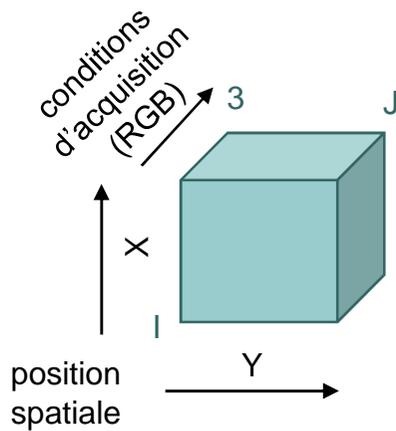


Images multivariées
ou hyperspectrales
(séquence d'images)



Types de données

- Exemples particuliers : Images multivariées ou hyperspectrales

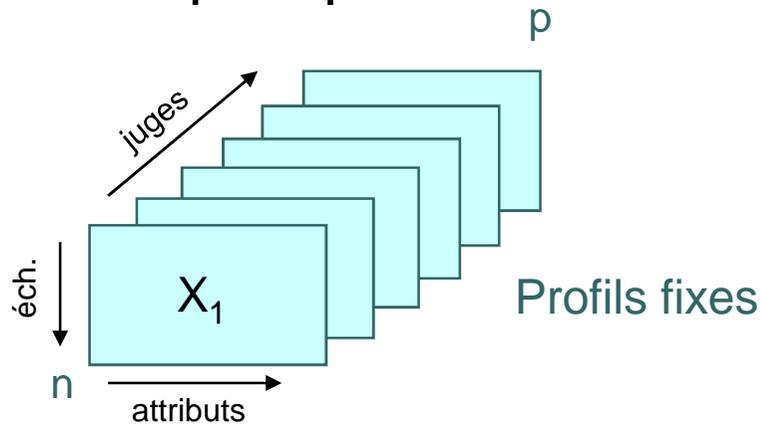


Multi-tableaux



Types de données

- Exemples particuliers : Données sensorielles (dégustation)

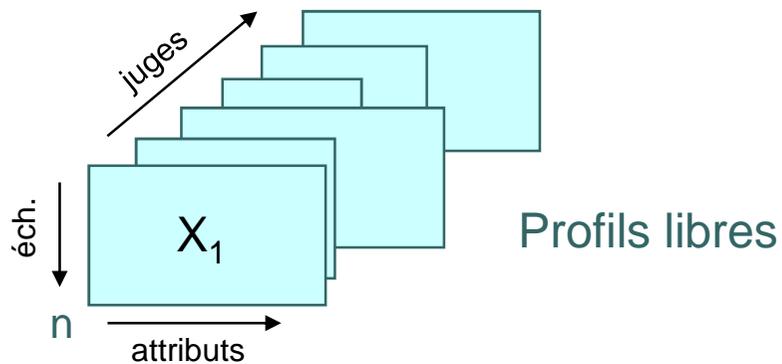


Objectif = s'affranchir de l'effet juge !



Multi-tableaux

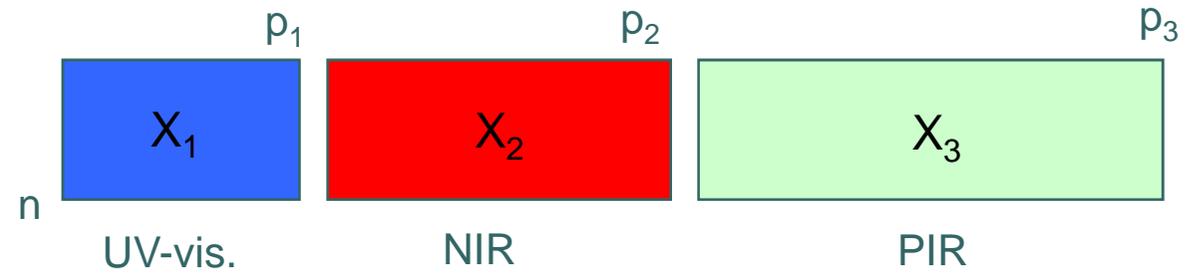
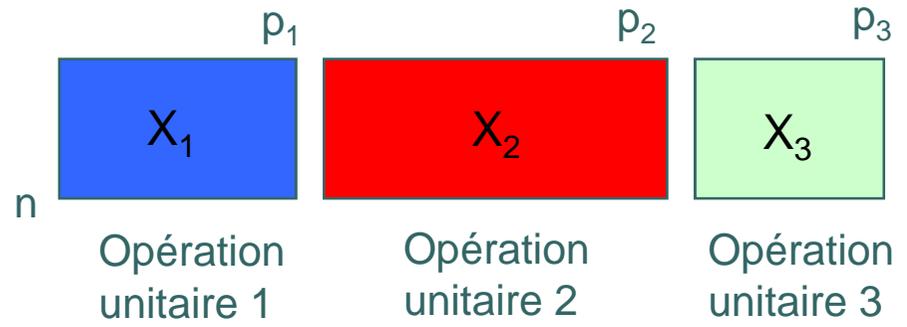
Pas multiway !



Multi-tableaux

Types de données

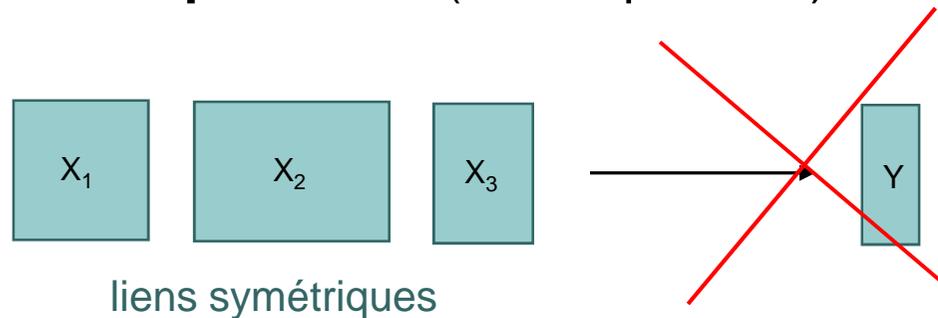
- Autres exemples de multi-tableaux :

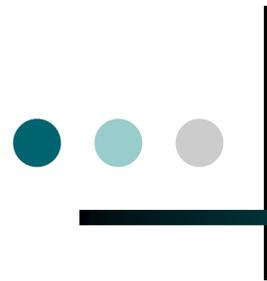


Objectifs

- Différents objectifs des analyses multi-tableaux :
 - Etude des liens entre différents types de mesures
 - Etude des complémentarités et redondances de différentes techniques analytiques
 - Etude du consensus d'un jury sensoriel
 - Identification de structures communes en imagerie
 - ...

Exploratoire (et non prédictif)



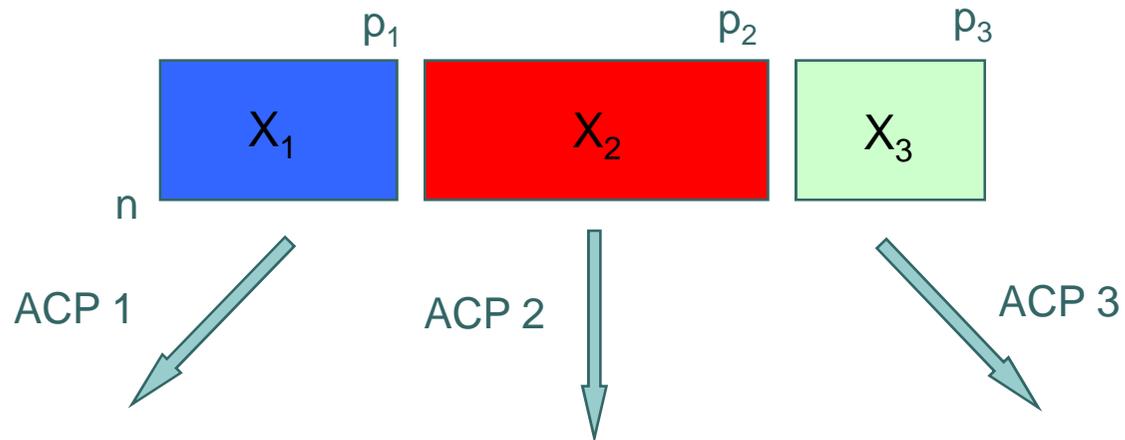


Sommaire

- I. Problématique
- II. Théorie**
- III. Exemple d'application
- IV. Conclusions

Principe

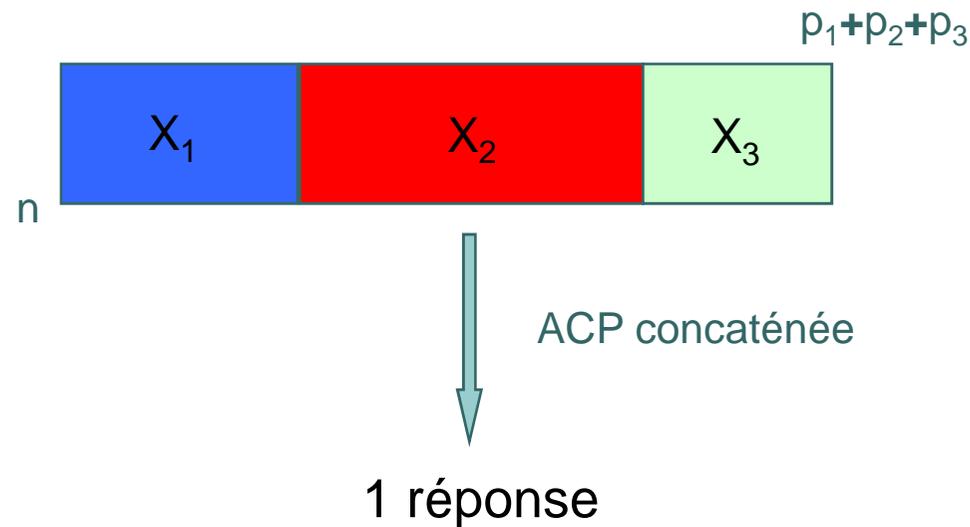
- Modélisation de la structure des données :



Plusieurs réponses ?

Principe

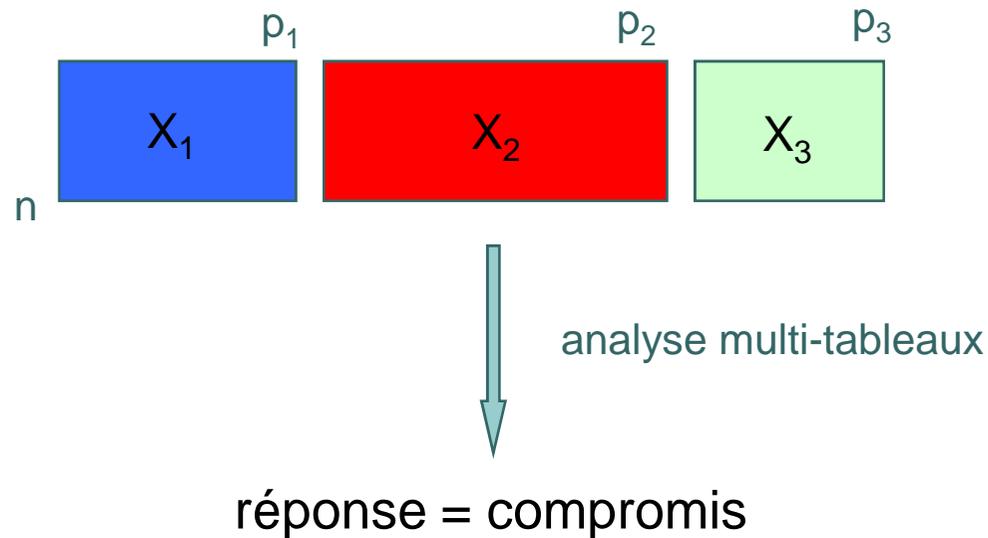
- Modélisation de la structure des données :



- Limites de l'ACP concaténée :
 - Poids des tableaux = f(nombre de variables, variabilité des mesures)
 - N'utilise pas explicitement le lien entre tableaux
 - Interprétation par tableau pas immédiate

Principe

- Solution = approche multi-tableaux :



- Comment construire le compromis ?
- Comment affecter un poids pour chaque bloc dans le compromis ?

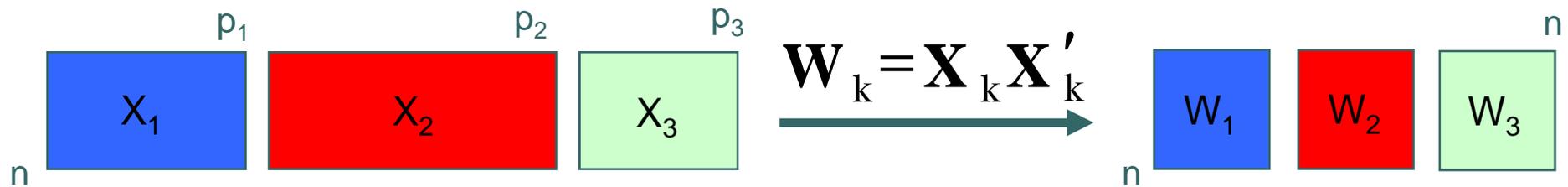


Plusieurs méthodes multi-tableaux pour des objectifs / données différents ...

● ● ● | **Théorie**

○ Analyse en Composantes Communes et Spécifiques (ACCPS) :

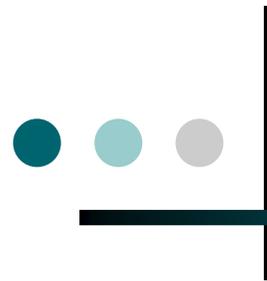
Qannari *et al.* (1997) – Nielsen *et al.* (2001) – Mazerolles *et al.* (2002) – Hanafi *et al.* (2004) - Preys *et al.* (2006)



$$\mathbf{W}_k = \sum_{j=1}^n \lambda_j^{(k)} \mathbf{q}_j \mathbf{q}'_j + \mathbf{E}_k$$

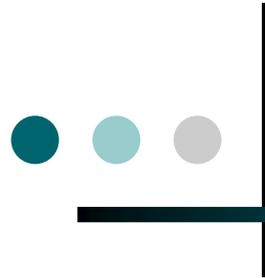
↙ poids spécifiques

- Chaque composante commune q_j est pondérée par un scalaire $\lambda_j^{(k)}$ reflétant la contribution de la table X_k à la composante q_j → structures communes et spécifiques



Sommaire

- I. Problématique
- II. Théorie
- III. Exemple d'application**
- IV. Conclusions



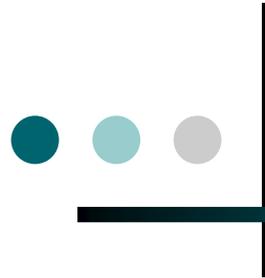
Optimisation d'une méthode analytique rapide par MS

○ Contexte :

- Méthodes de référence de caractérisation de la composition polyphénolique des vins rouges lentes et coûteuses
→ développement d'une méthode analytique rapide par injection directe en MS
- 2 modes d'ionisation et 2 modes de sélection d'ions sont testés

○ Objectif :

Déterminer à partir d'une campagne de mesures quel est (sont) le(s) mode(s) d'acquisition des spectres le plus pertinent pour cette analyse rapide ?



Matériel et méthodes

- Echantillons :
 - 62 vins rouges commerciaux (AOC Beaujolais)
 - Sources de variabilité de la composition polyphénolique :
 - Cépage (Gamay ou non)
 - Origine géographique (différentes AOC Beaujolais et non AOC)
 - Millésime (1999 à 2003)
- Injection directe en MS :
 - Sources APCI et ESI
 - Modes + et –
 - Analyseur TOF

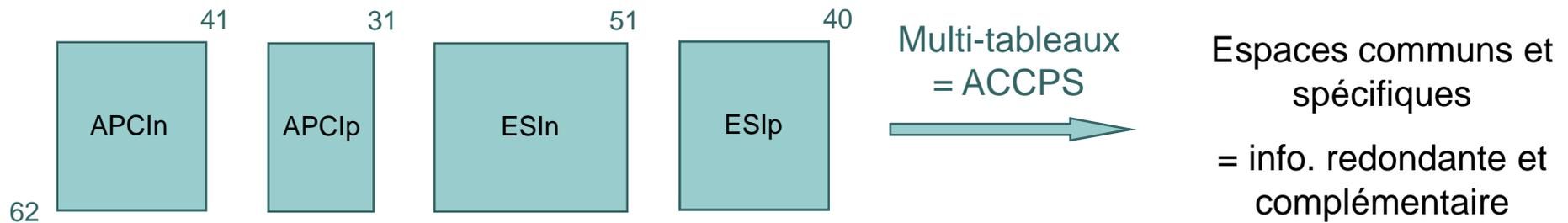
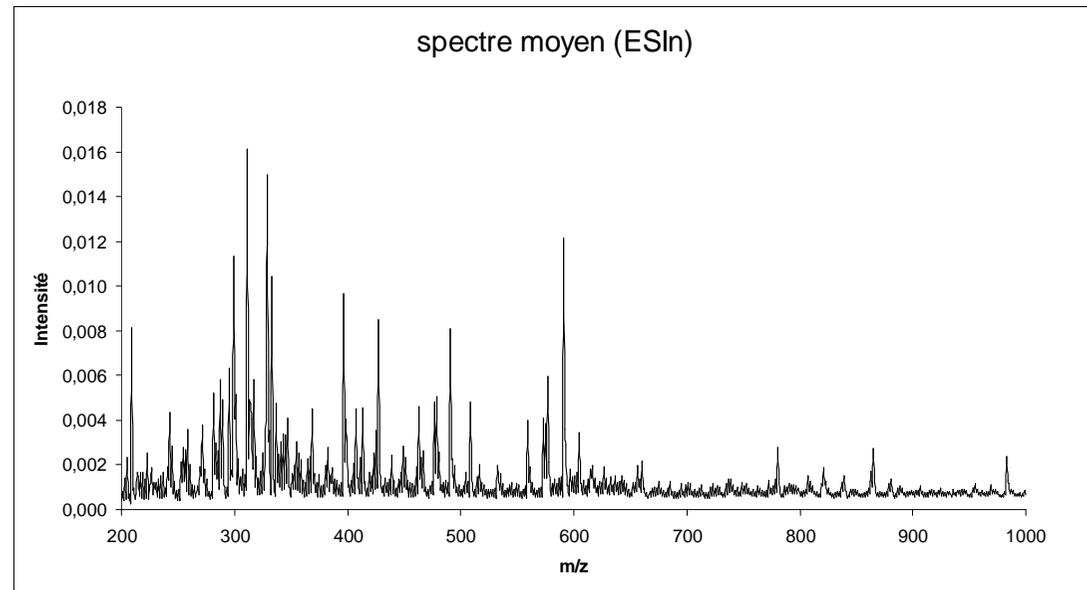
III. Exemple d'application



Chimiométrie

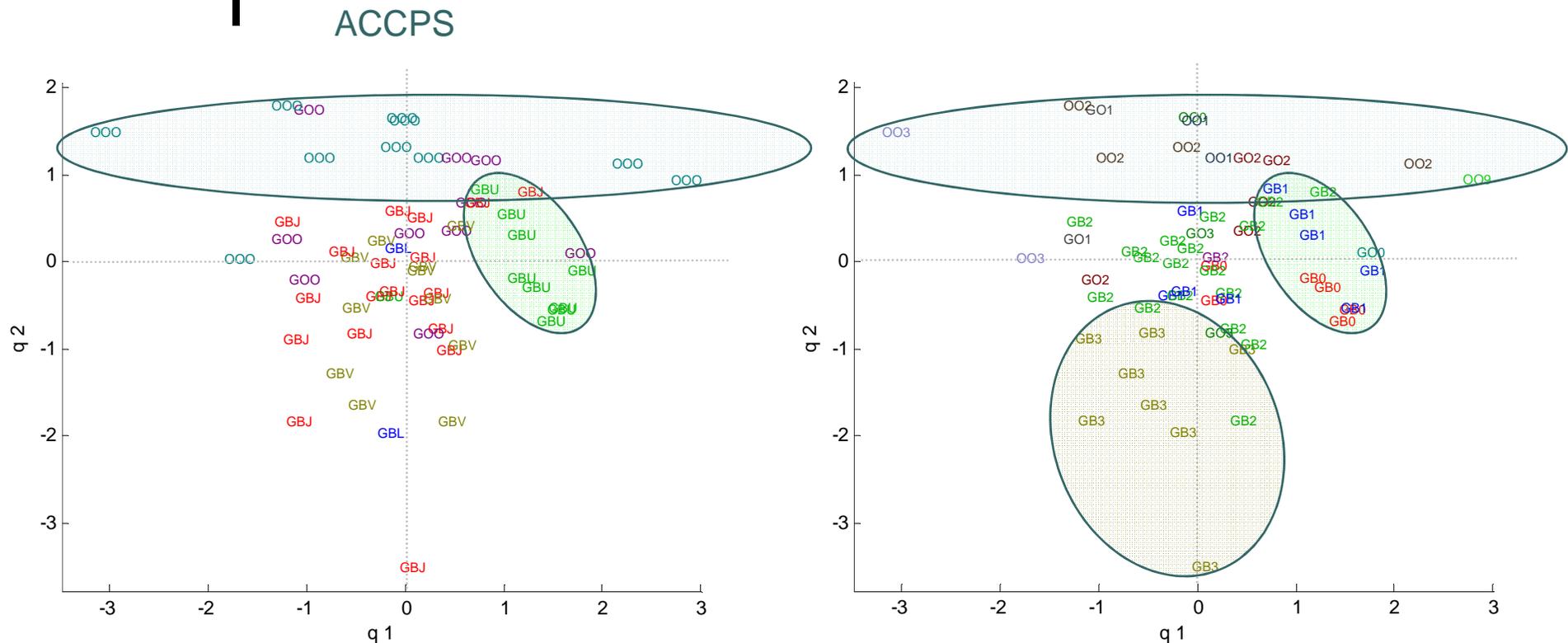
Pré-traitements :

- Recalage des masses
- Sélection de variables (↘ bruit)
- Normalisation en ligne



III. Exemple d'application

Résultats : choix des techniques d'acquisition



	q1	q2
APCIn	0.40	0.07
APCIp	0.65	0.05
ESIn	0.10	0.42
ESIp	0.72	0.02

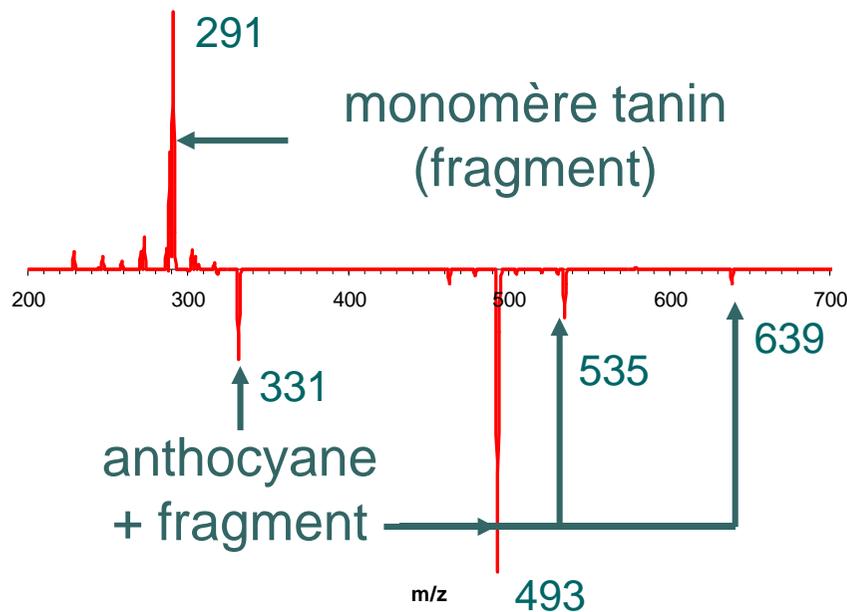


- Complémentarité des modes d'acquisition pour acquérir une info. utile
 - Redondance pour certains modes
- choix de combiner 2 modes

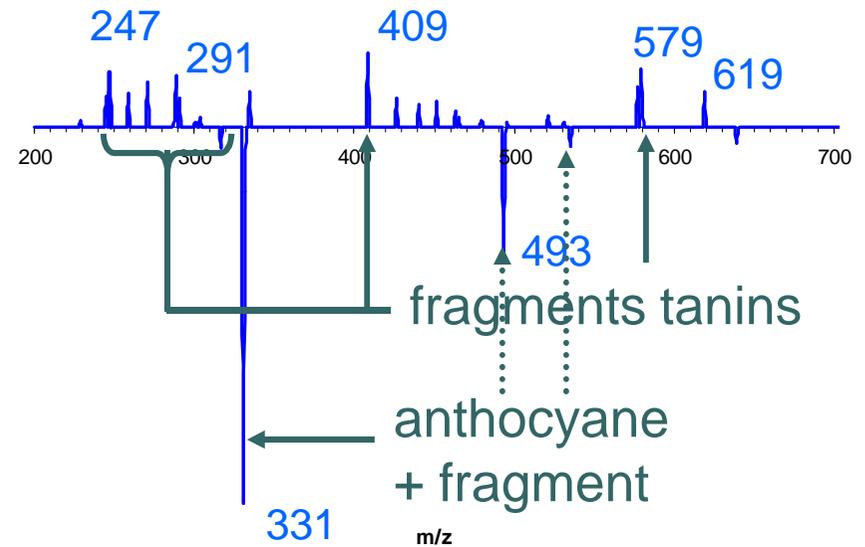
III. Exemple d'application



Résultats : outils d'interprétation



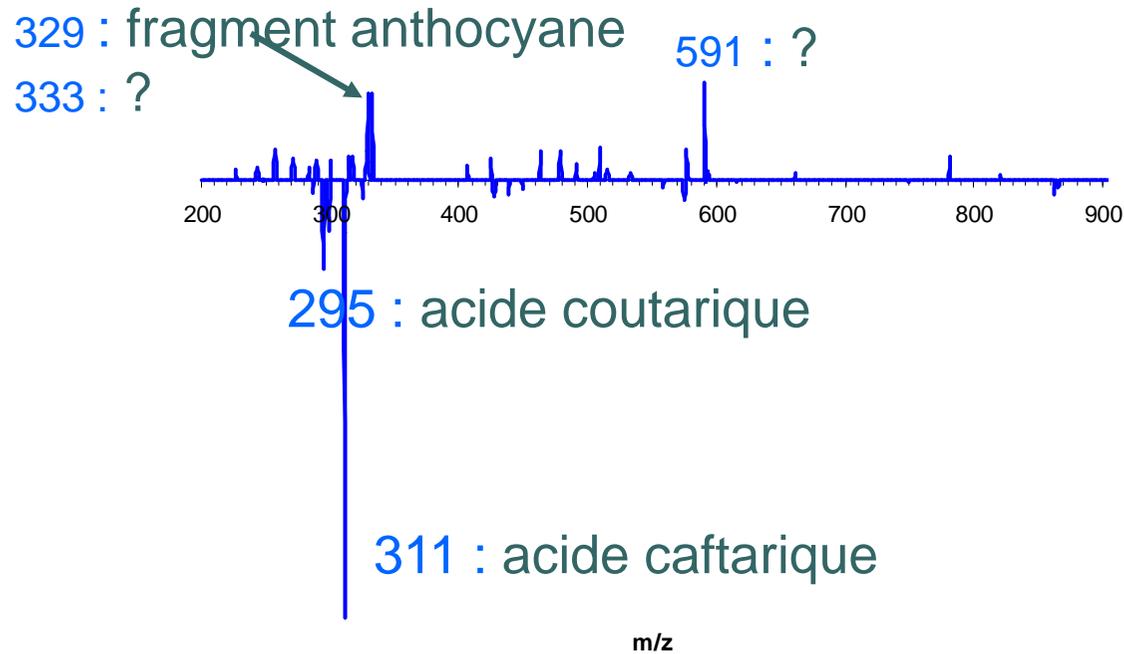
Loadings sur q1 pour le tableau APCIp



Loadings sur q1 pour le tableau ESIp

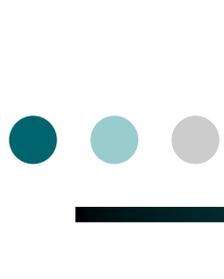


Résultats : outils d'interprétation

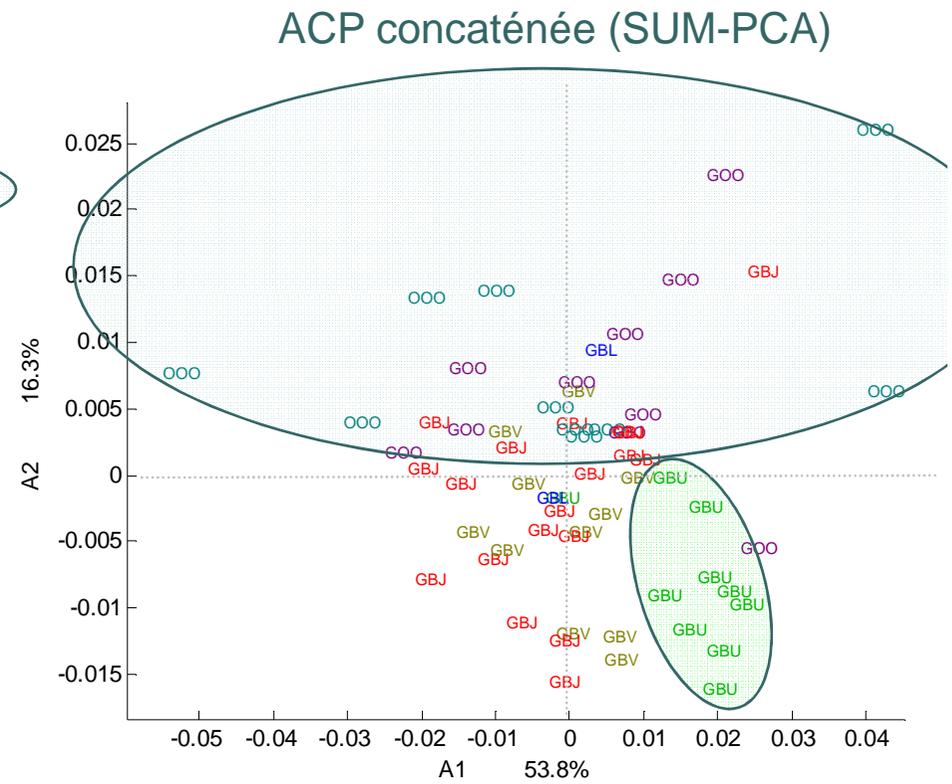
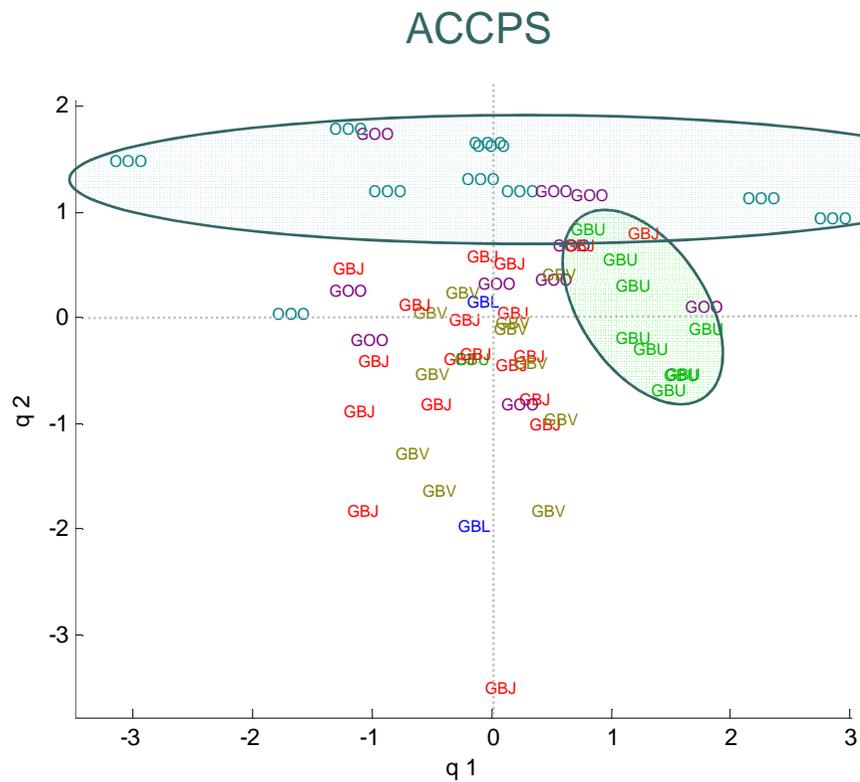


Loadings sur q2 pour le tableau ESIn

III. Exemple d'application



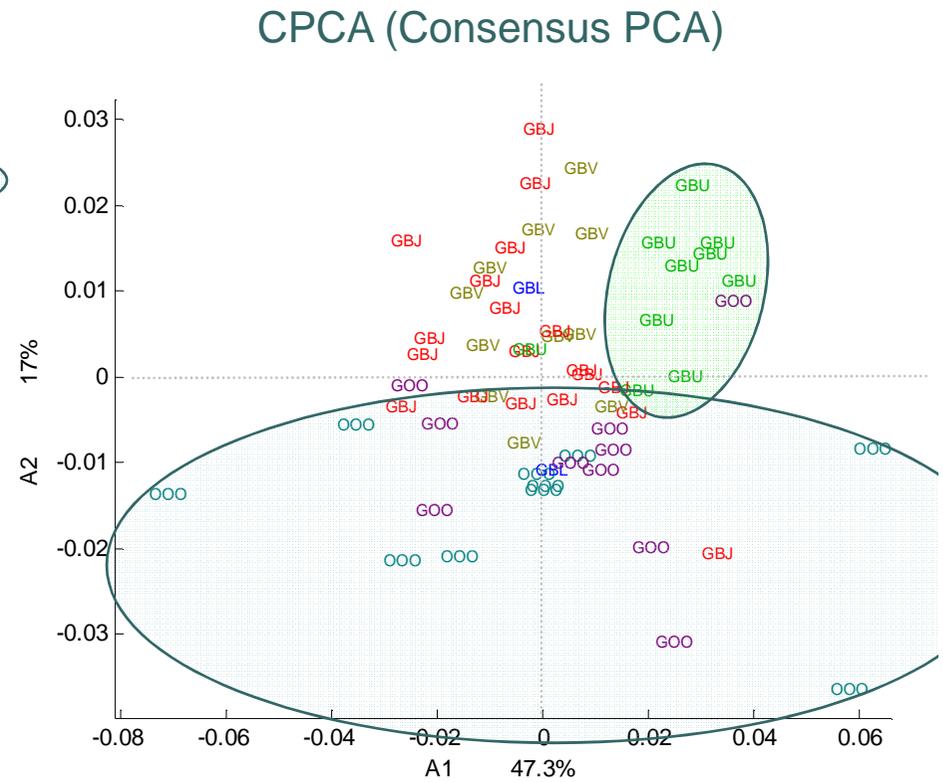
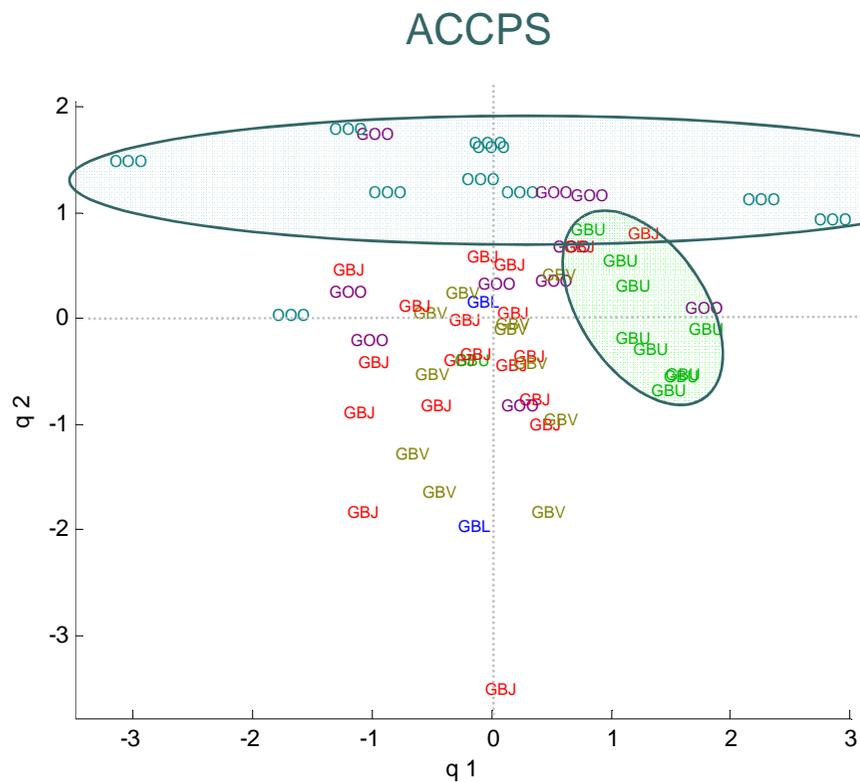
Comparaison avec l'ACP



III. Exemple d'application

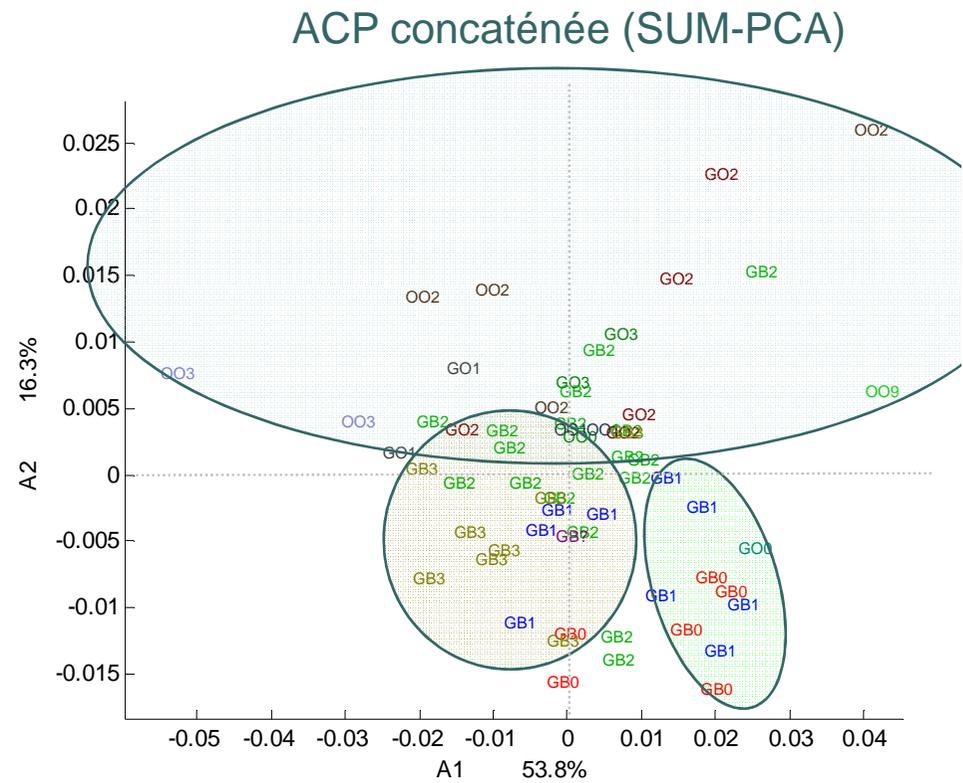
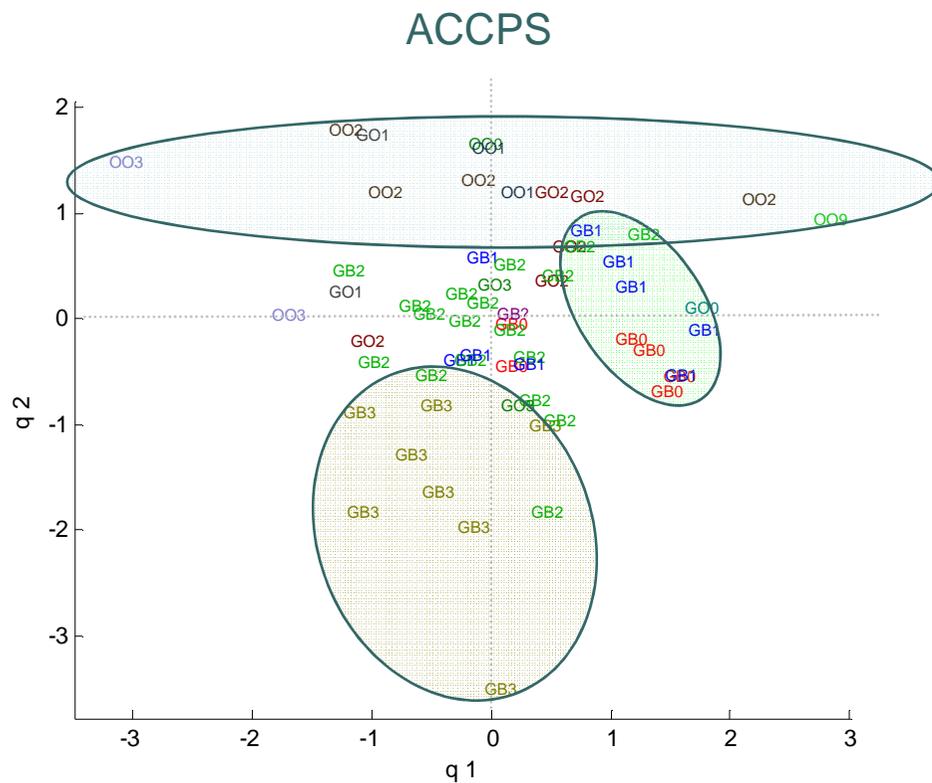


Comparaison avec l'ACP





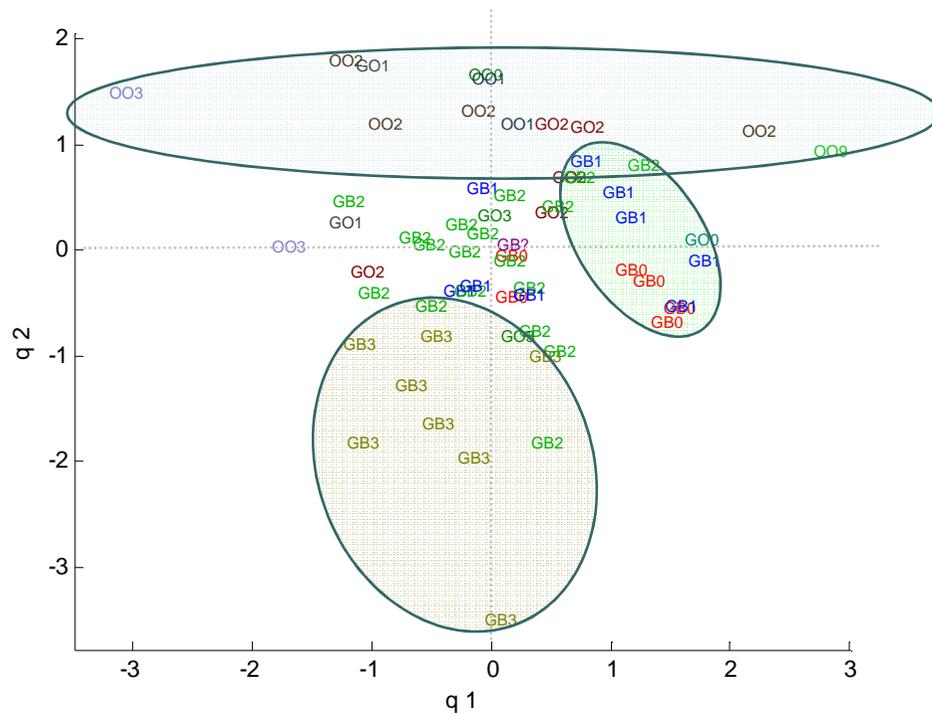
Comparaison avec l'ACP



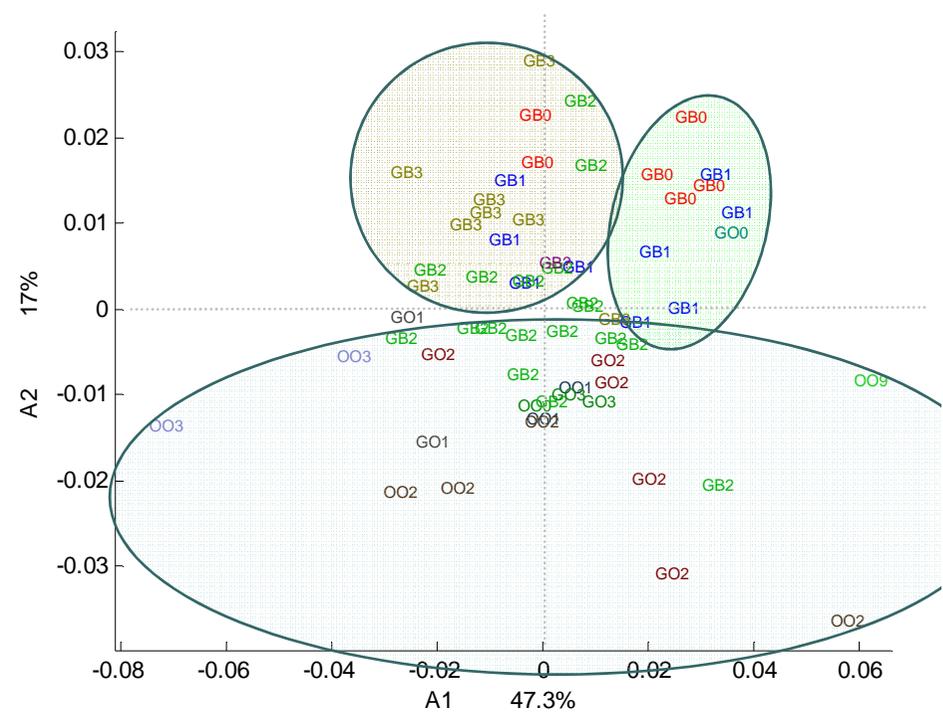


Comparaison avec l'ACP

ACCPS

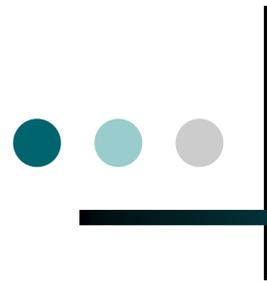


CPCA (Consensus PCA)



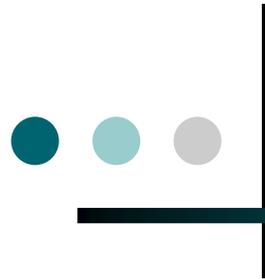
Multi-tableaux = utilisation lien entre blocs :

→ Phénomènes sous-jacents mieux modélisés (moins de bruit)



Sommaire

- I. Problématique
- II. Théorie et différentes approches
- III. Exemple d'application
- IV. Conclusions**



Conclusions

- Approche multi-tableaux (*multiblock*) utile :
 - Caractériser les échantillons mesurés par plusieurs techniques en utilisant le lien entre tableaux
 - Interprétation par rapport à chacun des tableaux
- Type de données et objectifs déterminent le choix de l'outil multi-tableaux :
 - Recherche de l'info. commune et spécifique → ACCPS, ACOM, ...
 - Recherche d'un consensus (sensoriel) → STATIS, AFM, ...
- De nombreuses applications sont encore à explorer !...