

Doctorat de l'Université Louis Pasteur

Imagerie chimique pour l'analyse des formes pharmaceutiques solides

Christelle Gendrin, 13 novembre 2008

F. Hoffmann-La Roche, PTGF-Q PRS, Bâle, Suisse

*Laboratoire des Sciences de l'Image de l'Informatique et de la Télédétection
LSIIT UMR-CNRS 7005, Université de Strasbourg*



Pourquoi cette thèse?

- Laboratoire de spectroscopie PIR : 2002
- Achat d'un système d'imagerie PIR : février 2005
- Thèse débutée en décembre 2005
- Développer et tester des méthodes de traitement des données d'imagerie pour les applications pharmaceutiques

Introduction

Distribution des composés

Détermination de la teneur en principe actif

Conclusions

Introduction

L'initiative PAT

PAT: Process Analytical Technology
QbD: Quality by Design

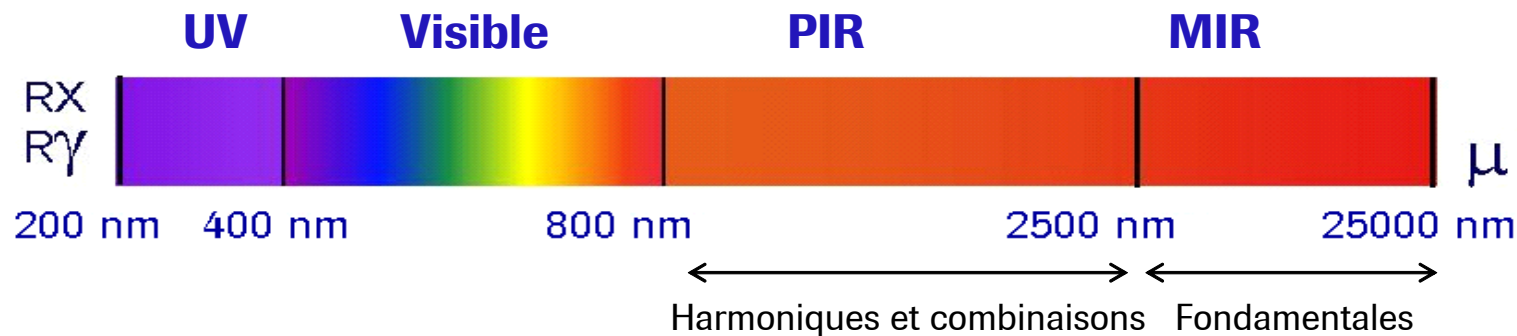
Promouvoir la mise en place de systèmes permettant une analyse continue des paramètres critiques du procédé

- Compréhension du procédé
- Implémentation de nouvelles technologies
- A terme : contrôle qualité en ligne

Introduction

La spectroscopie Proche Infrarouge

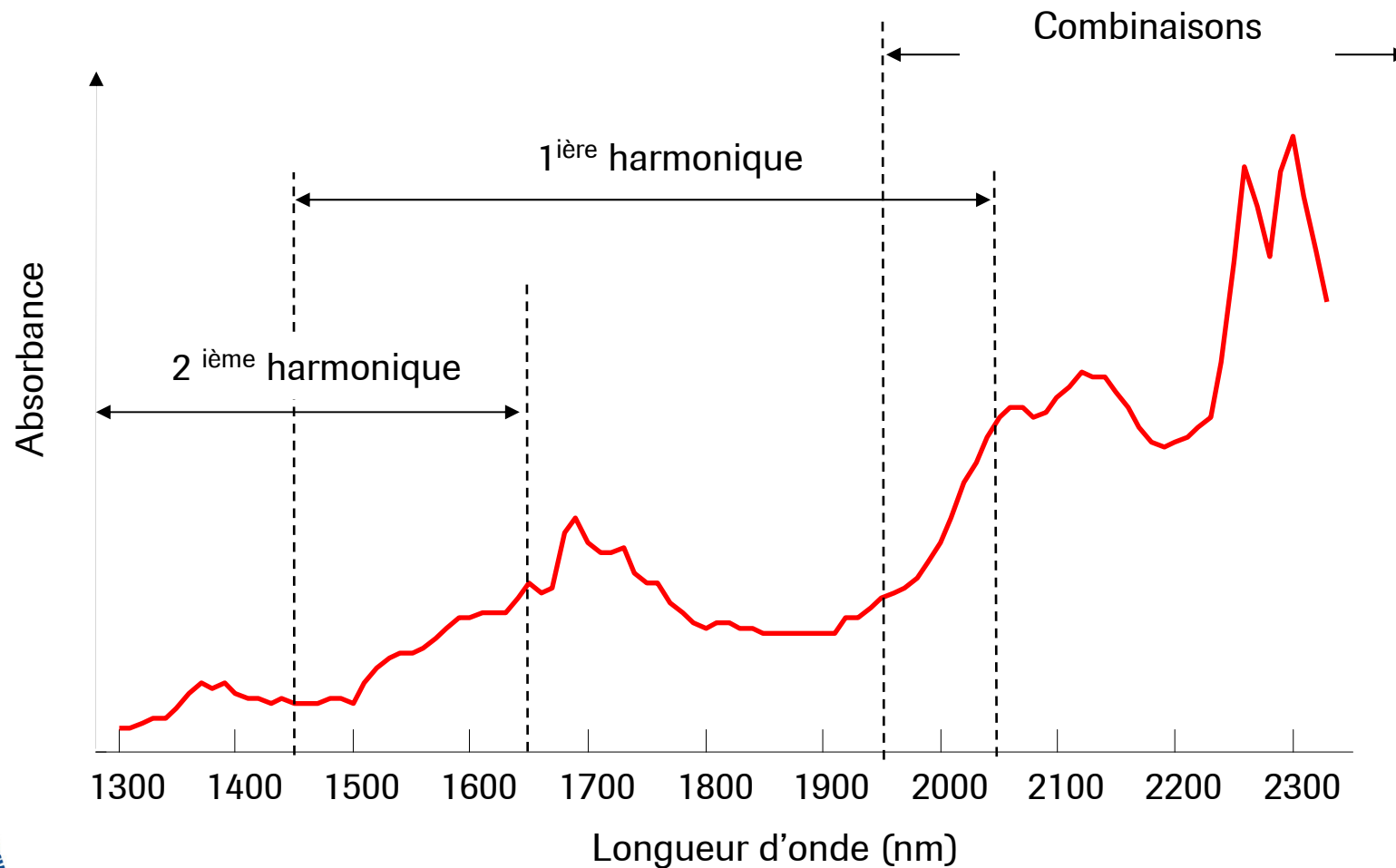
- Spectroscopie vibrationnelle



- Avantages :
 - Non-destructif
 - Rapide
 - Pas de préparation des échantillons
 - Les spectres incluent informations physique et chimique
 - Analyse à travers le verre est possible

Introduction

La spectroscopie Proche Infrarouge



Introduction

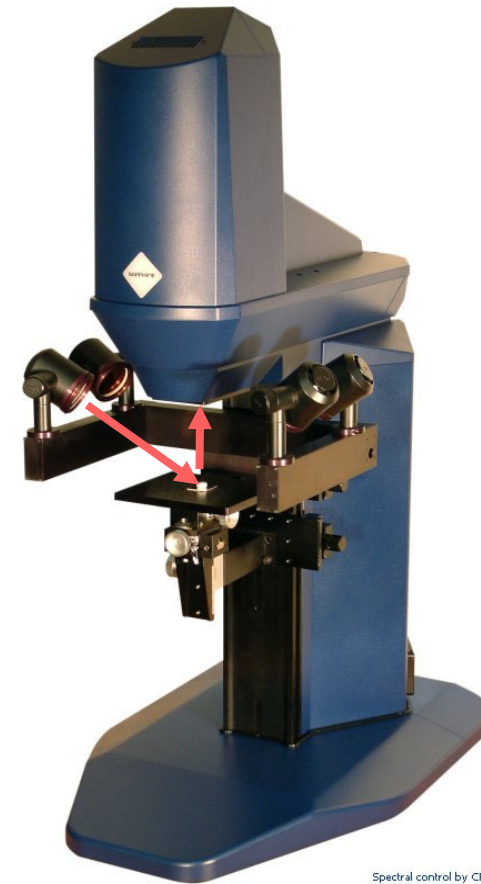
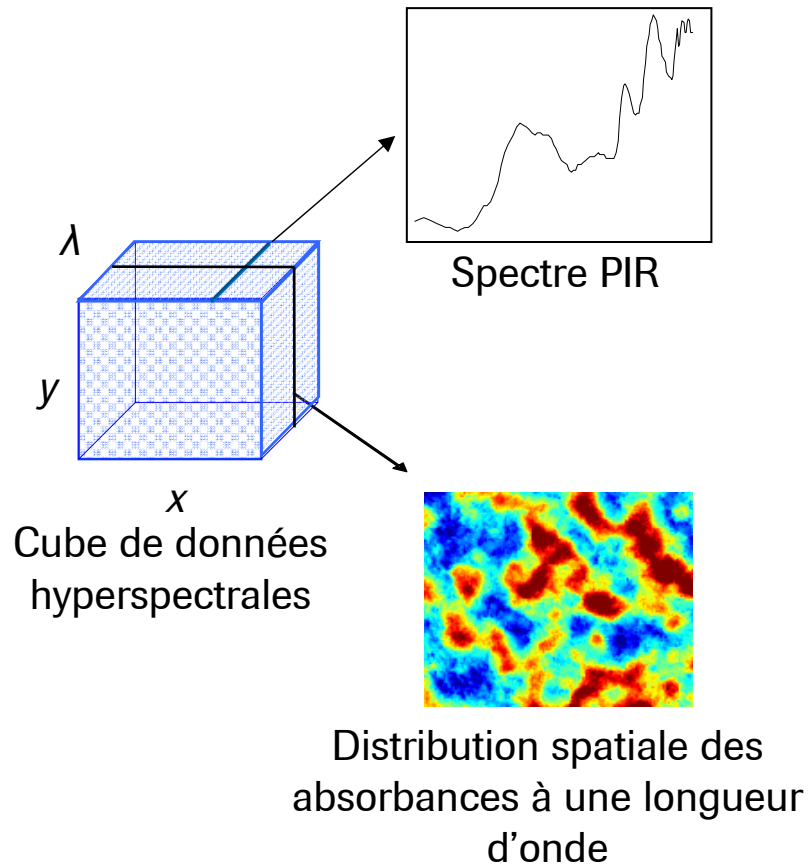
Les applications de la PIR classique

- Contrôle des procédés :
 - Teneur en eau
 - Teneur en Principe Actif (PA)
 - Identification des poudres avant la production
 - Fin du temps de réaction, homogénéité des mélanges

- Au laboratoire :
 - Identification des contrefaçons

Introduction

L'imagerie Proche Infrarouge



Introduction

Objectifs de la thèse

Extraire les cartes de distribution des composés et
estimer la teneur en PA par imagerie chimique
Proche Infrarouge

Introduction

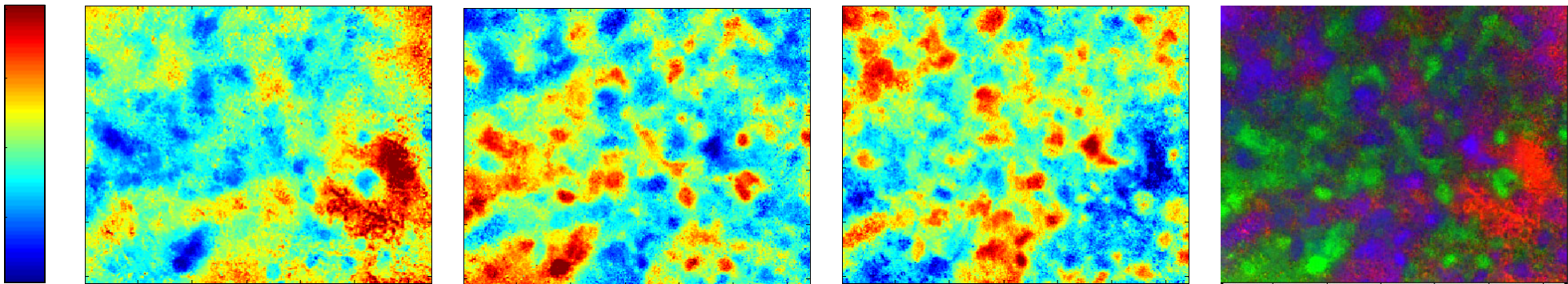
Distribution des composés

Détermination de la teneur en principe actif

Conclusions

Distribution des composés

Analyse univariée : analyse à une longueur d'onde



1680 nm : principe actif

2260 nm : povidone

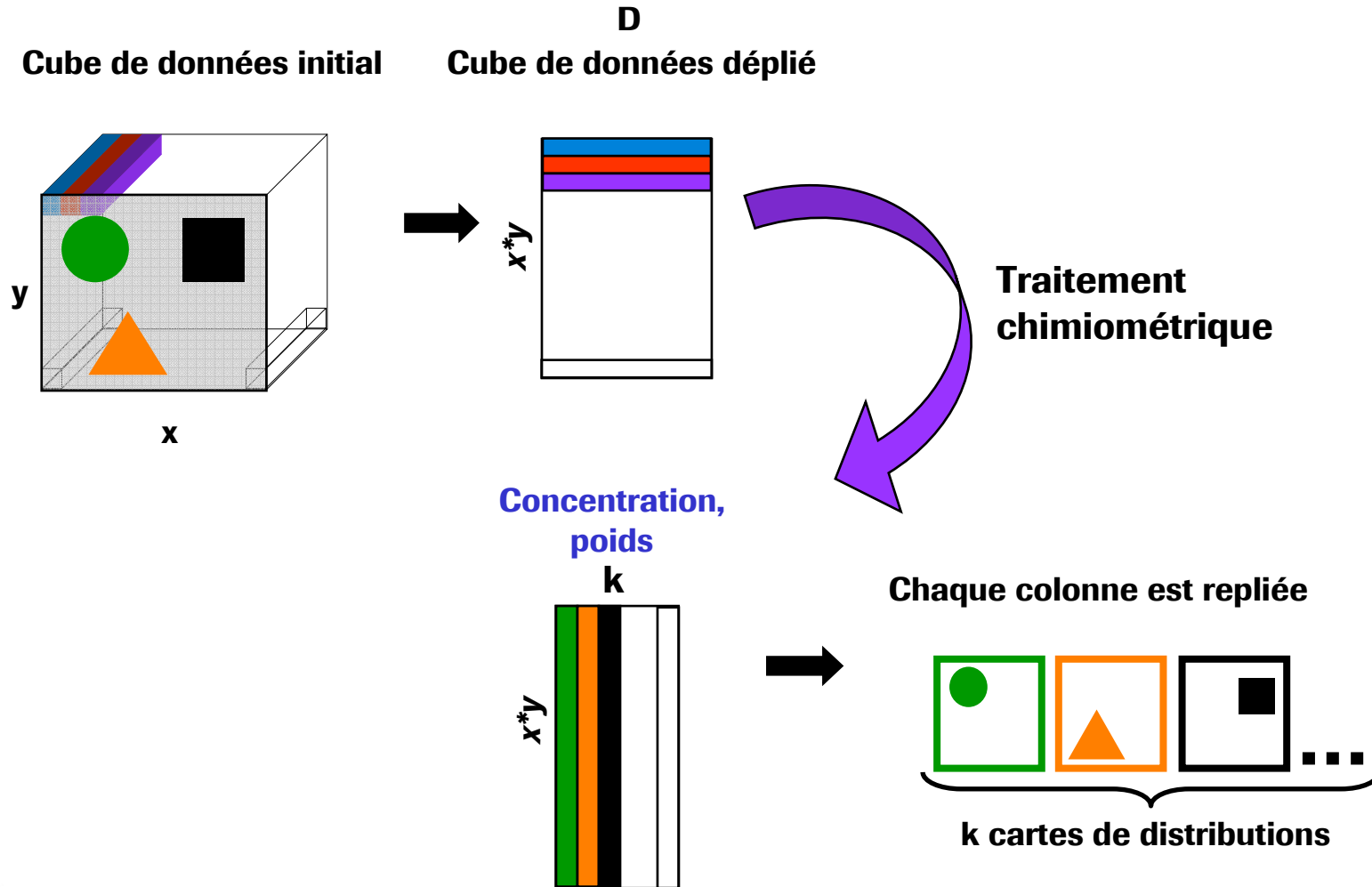
2100 nm : amidon

Image RVB

- Avantage : simple
- Inconvénient : superposition des bandes d'absorbances

Distribution des composés

Analyse multivariée : traitement des données

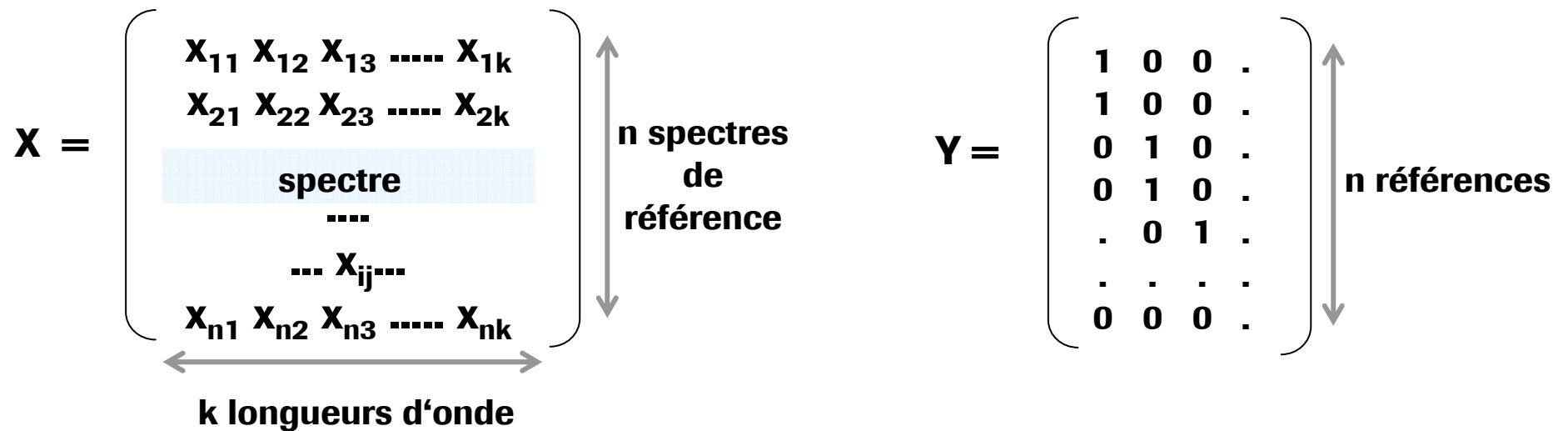


Distribution des composés

Analyse multivariée avec spectres de référence : PLS-DA

- PLS-DA : Partial Least Squares - Discriminant Analysis
- Une librairie spectrale est utilisée pour faire un modèle mathématique

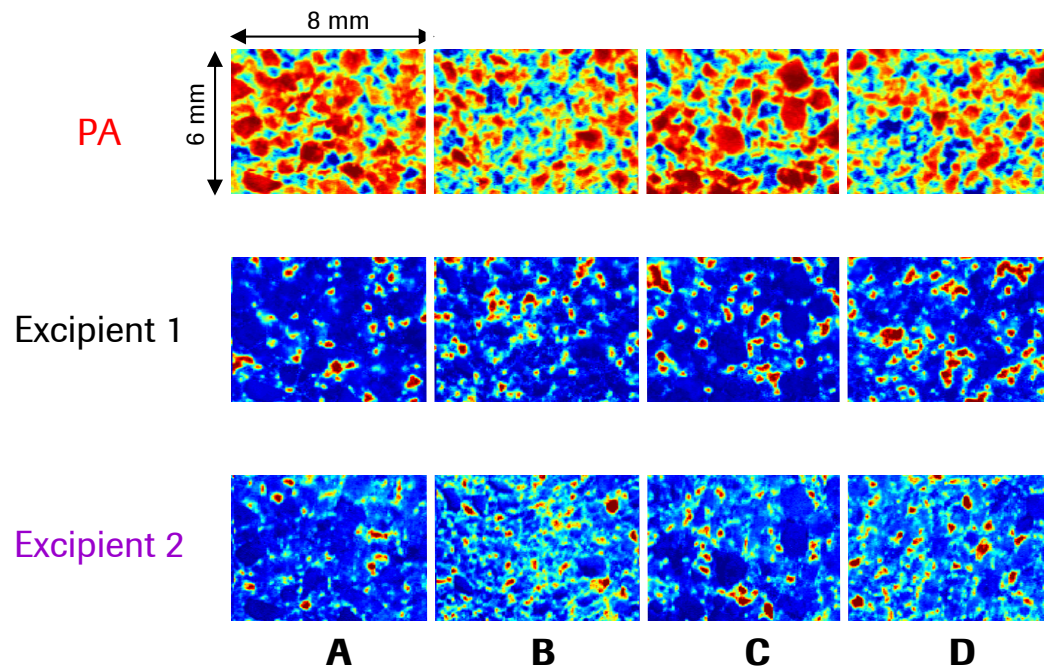
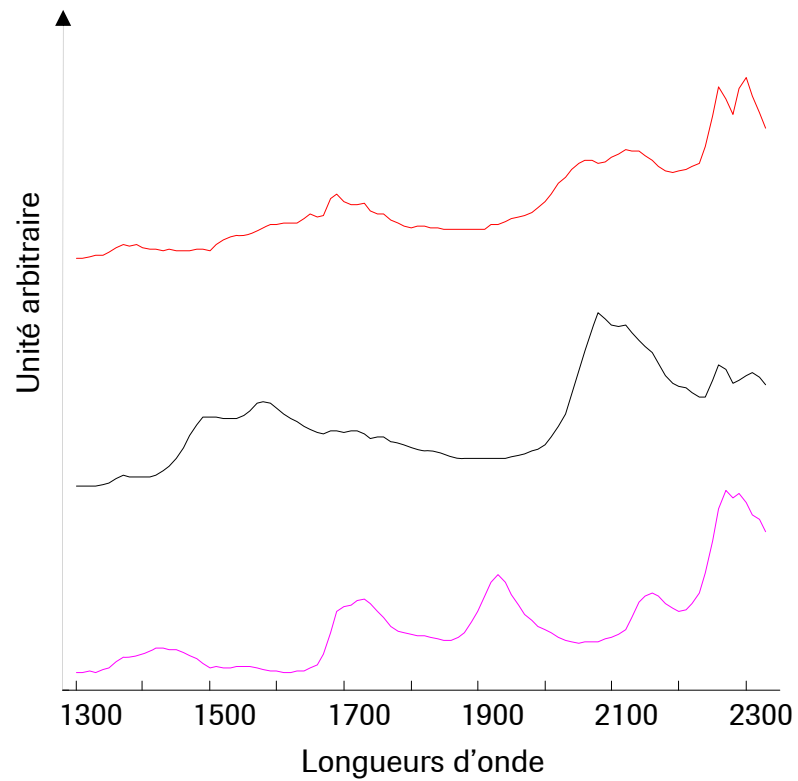
x_{ij} absorbance



Distribution des composés

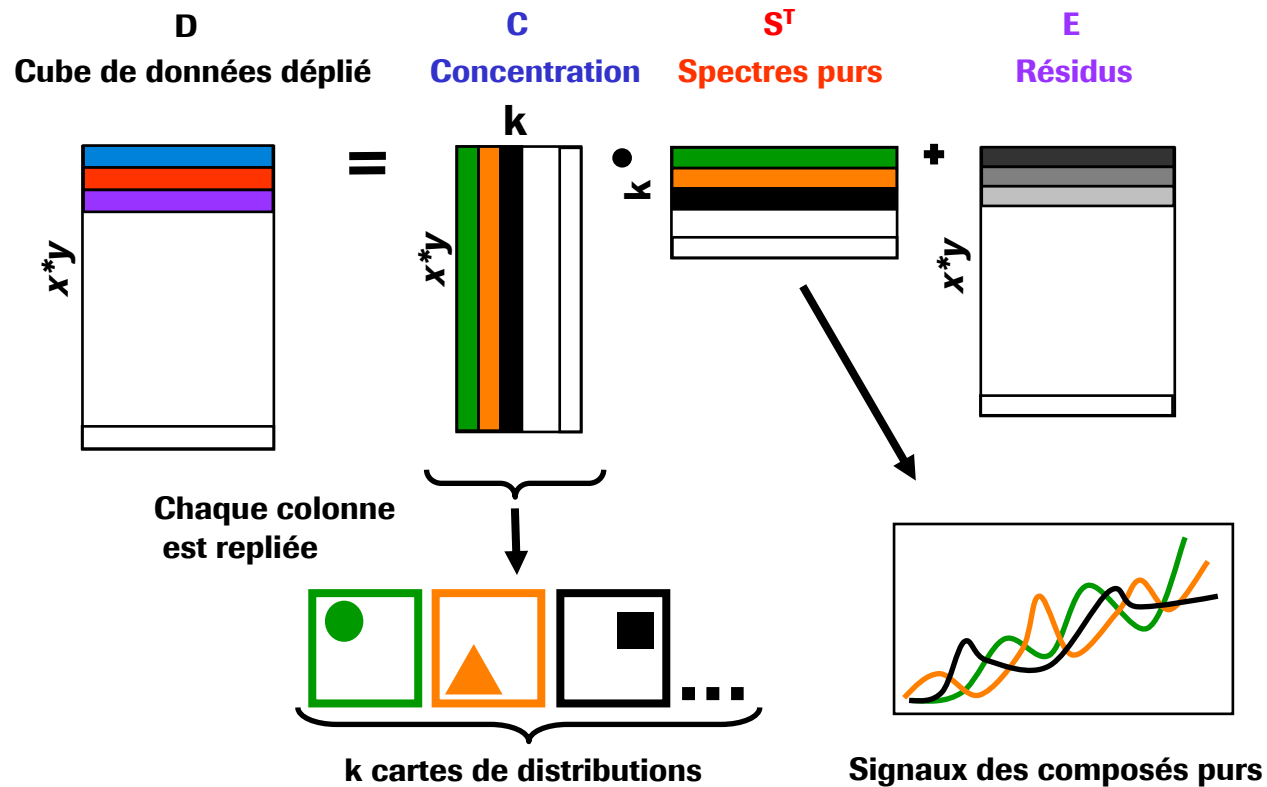
Analyse multivariée avec spectres de référence : PLS-DA

- PLS-DA



Distribution des composés

Sans a priori : résolution de courbes multivariées



- Solutions multiples
- Des contraintes sont imposées : ici contraintes de positivité

Distribution des composés

Sans a priori : algorithmes

- MCR-ALS : *Multivariate Curve Resolution – Alternating Least Squares*

$$\| \mathbf{D} - \mathbf{C}\mathbf{S}^T \|^2$$

Sous les contraintes : $\mathbf{C} > 0$ and $\mathbf{S}^T > 0$

Estimation initiale de \mathbf{S}_0 ou \mathbf{C}_0

Itération:

(1) Avec \mathbf{D} et \mathbf{S}_i^c calcul de \mathbf{C}_{i+1} (moindres carrés)

Contraindre $\mathbf{C}_{i+1} \rightarrow \mathbf{C}_{i+1}^c$

(2) Avec \mathbf{D} et \mathbf{C}_{i+1}^c calcul de \mathbf{S}_{i+1} (moindres carrés)

Contraindre $\mathbf{S}_{i+1} \rightarrow \mathbf{S}_{i+1}^c$

Distribution des composés

Sans a priori : algorithmes

- PMF : *Positive Matrix Factorization*
- Le critère suivant est minimisé

$$F(\mathbf{E}) = \sum_{i,j} \frac{(\mathbf{D} - \mathbf{C}\mathbf{S}^T)_{i,j}^2}{\sigma_{i,j}^2}$$

Sous les contraintes : $\mathbf{C} > 0$ and $\mathbf{S}^T > 0$

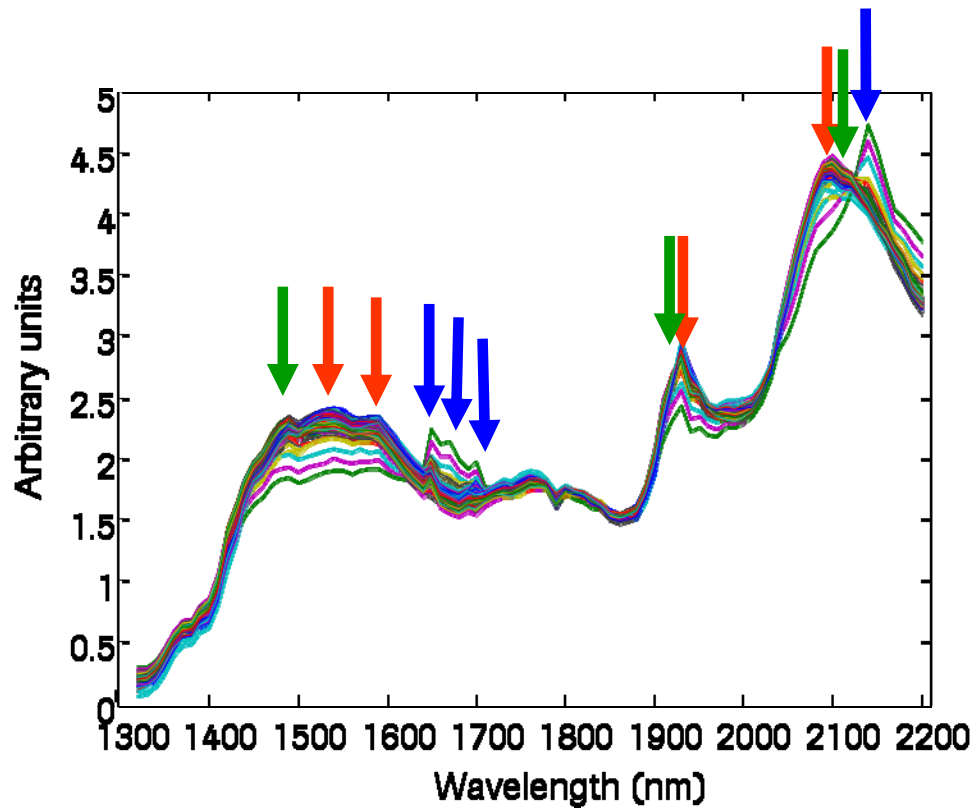
ou $\sigma_{i,j}$ désigne l'incertitude sur la donnée (i,j),

- La mesure est effectuée cinq fois afin d'estimer les incertitudes
- Multilinear Engine (ME) est utilisé pour l'optimisation

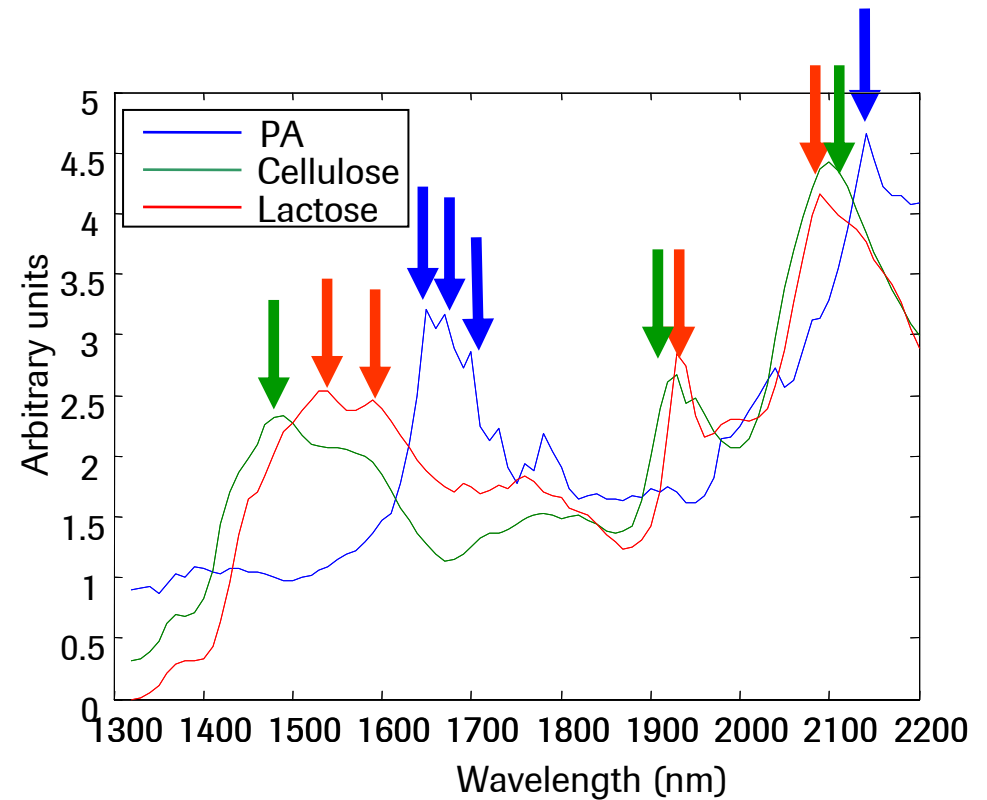
Distribution des composés

Sans a priori : données

- Lot : actif (5%), cellulose (50%), lactose (45%)



Spectres mélangés (matrice D)



Spectres de références

Distribution des composés

Sans a priori : évaluation de l'estimation

- Corrélation

$$\text{Corr}_k = \frac{\sum_j (s_{kj} - \bar{s}_k)(\text{ref}_{kj} - \overline{\text{ref}}_k)}{\sqrt{\sum_j (s_{kj} - \bar{s}_k)^2} \sqrt{\sum_j (\text{ref}_{kj} - \overline{\text{ref}}_k)^2}}$$

- RSD: Rapport Signal à Distorsion

Saïd Moussaoui et al, IEEE transactions on signal processing, Vol 54, 797-807, 1998

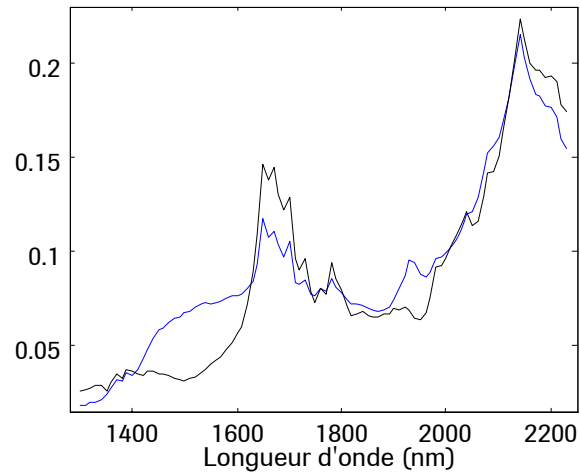
$$\text{RSD}_k = 10 \log_{10} \left(\frac{\sum_j \text{ref}_{kj}^2}{\sum_j (\text{ref}_{kj} - s_{kj})^2} \right)$$

- Concentration extraite

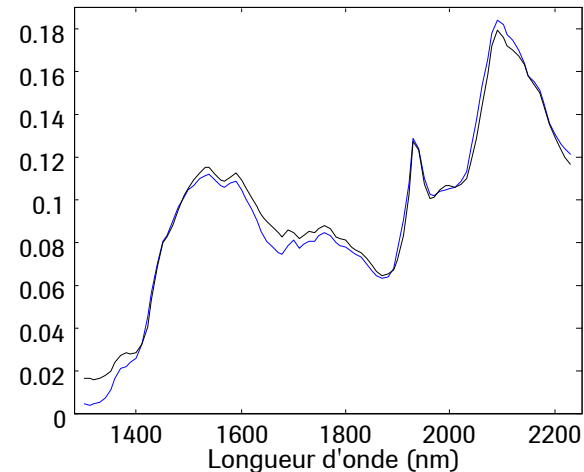
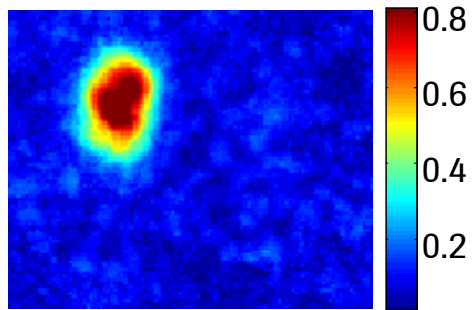
Distribution des composés

Sans a priori : résultats

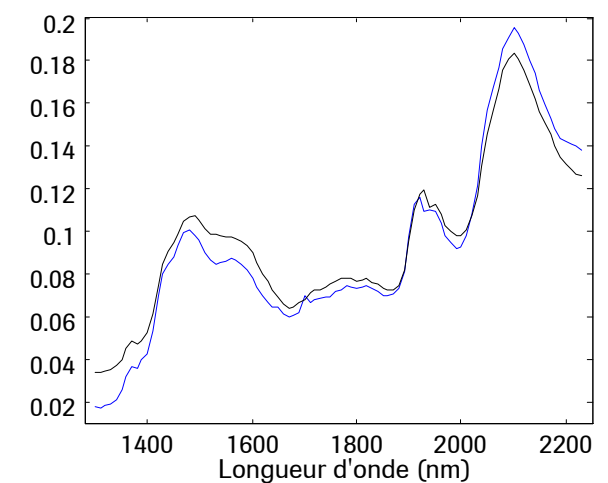
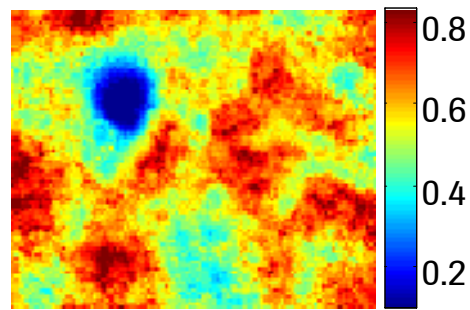
- MCR-ALS



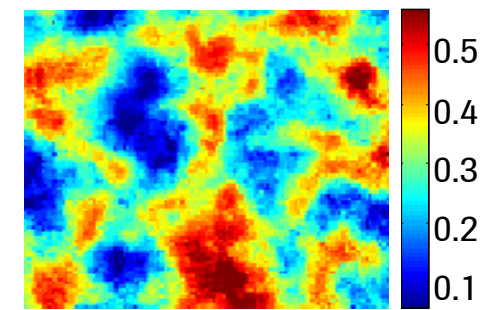
Facteur 1 = Actif



Facteur 2 = lactose



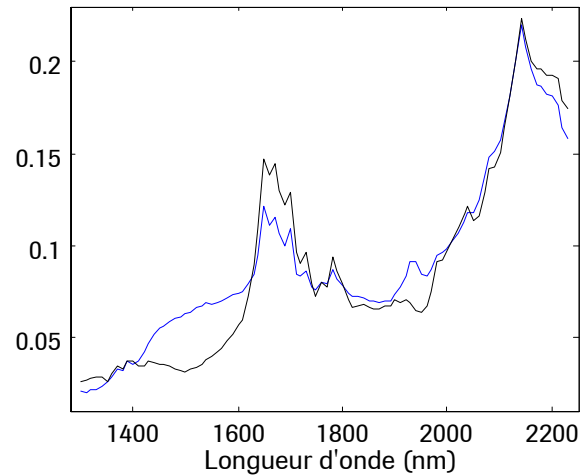
Facteur 3 = cellulose



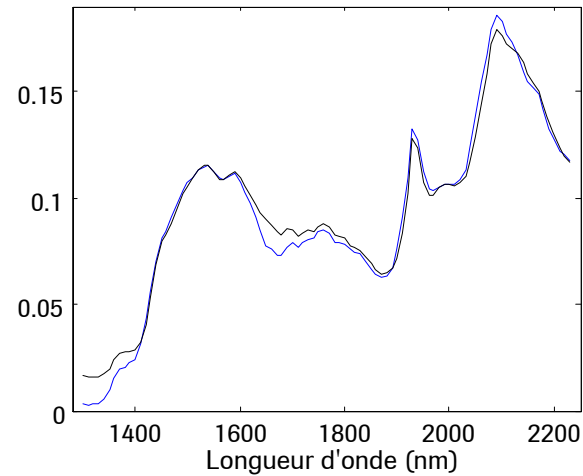
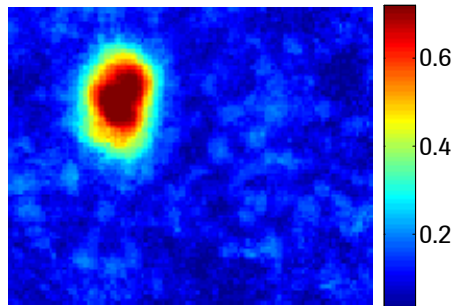
Distribution des composés

Sans a priori : résultats

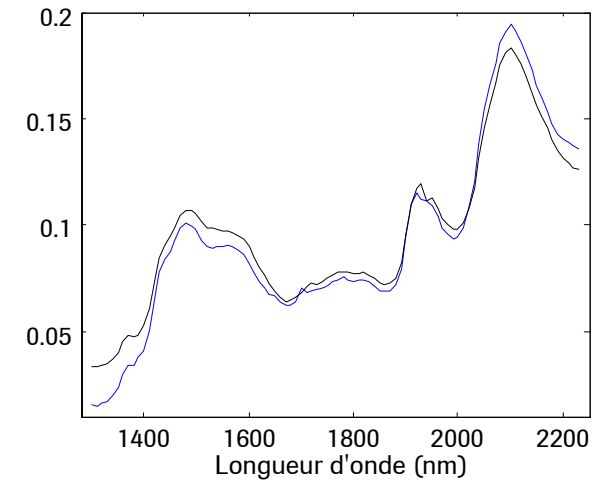
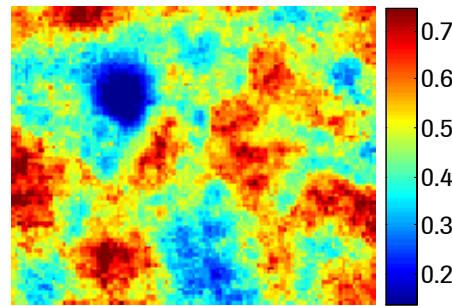
- PMF



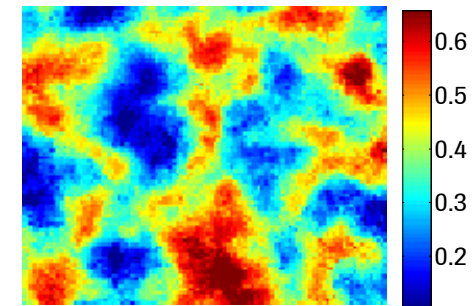
Facteur 1 = actif



Facteur 2 = lactose



Facteur 3 = cellulose



Distribution des composés

Sans a priori : résultats

		Corrélation	RSD	Concentration (%)
MCR-ALS	PA	0.954	15.5	12.2
	Cellulose	0.993	21.3	30.7
	Lactose	0.996	26.2	57.03
PMF	PA	0.968	16.8	14.2
	Cellulose	0.996	22.04	37.6
	Lactose	0.995	25.41	48.16

concentration du lot : actif (5%), cellulose (50%), lactose (45%)

Distribution des composés

Exploration des solutions

- Problèmes d'ambigüités rotationnelles
- **Paatero et al. Chemom. Intell. Lab. Syst., 60, 2002, 253-264.**
- **Paatero and Hopke, J. of Chemometrics, 2008, DOI: 10.1002/cem.1197**

- Améliorer l'extraction par :
 - Matrices de rotation

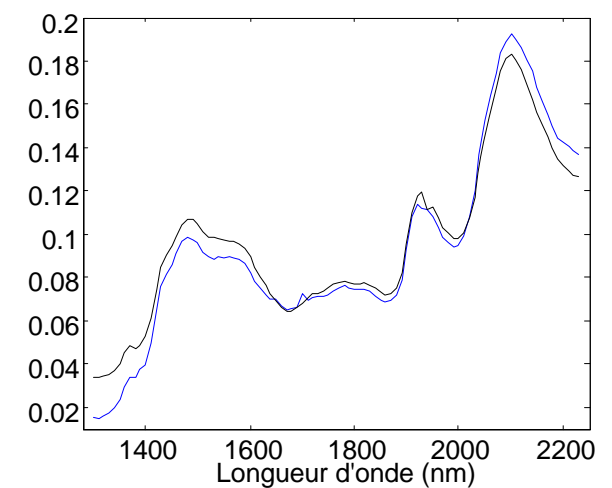
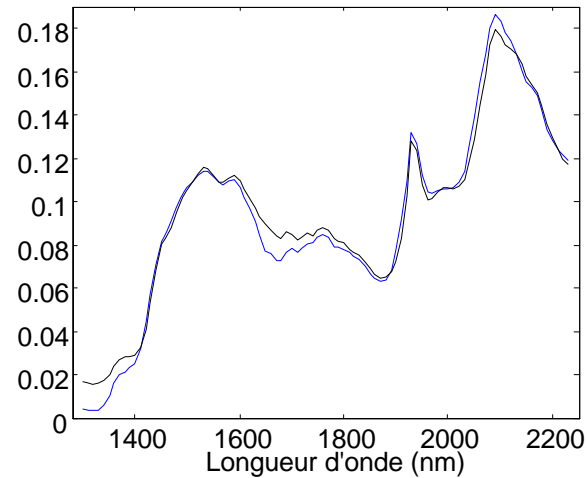
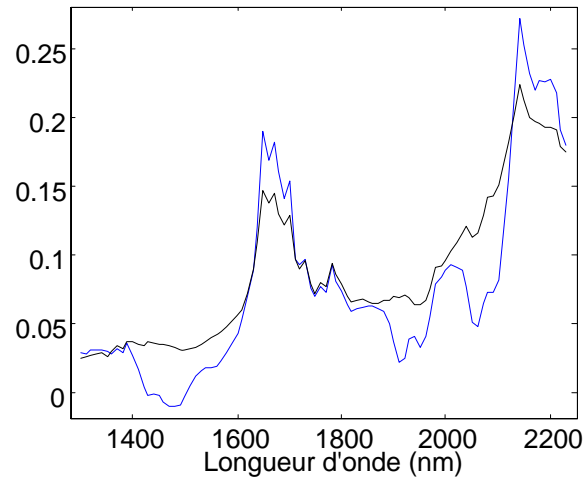
	C1'	C2'	C3'
C1	1	r	r
C2	-r	1	0
C3	-r	0	1

- Connaissances a priori : valeurs de concentration en un pixel, ou concentration moyennes

Distribution des composés

Exploration des solutions

- Matrice de rotation

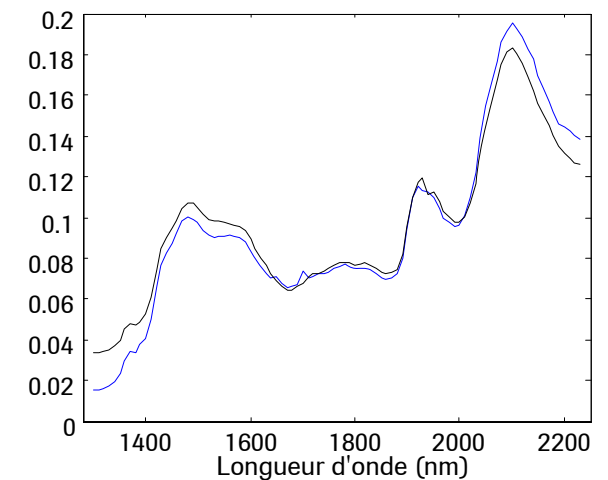
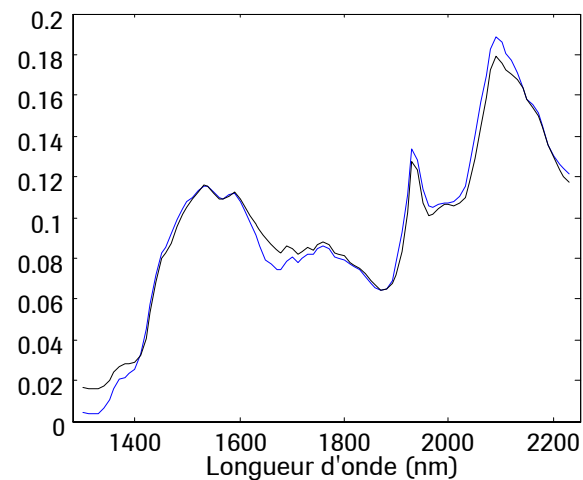
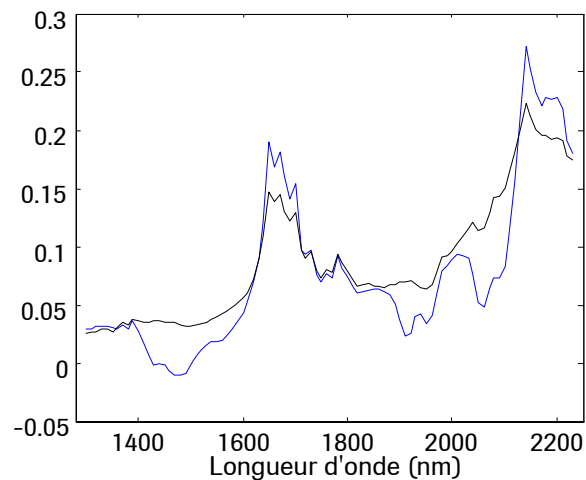


	Avant rotation			Après rotation		
	Corrélation	RSD	Conc. (%)	Corrélation	RSD	Conc. (%)
API	0.968	16.8	14.2	0.933	11.1	6.91
Cellulose	0.996	22.04	37.6	0.993	21.7	46.07
Lactose	0.995	25.41	48.16	0.995	25.3	48.43

Distribution des composés

Exploration des solutions

- Concentrations moyennes

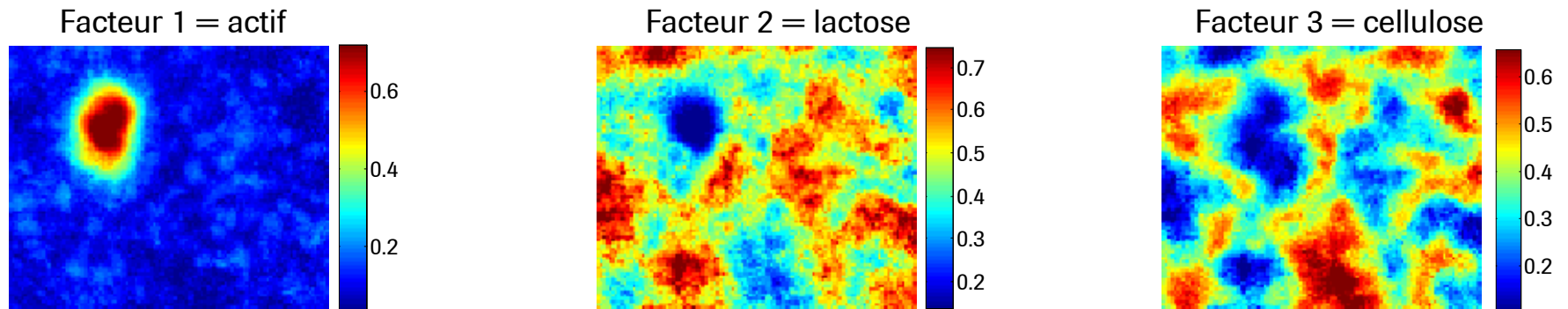


	Avant introduction des concentrations			Après introduction des concentrations		
	Corrélation	RSD	Conc. (%)	Corrélation	RSD	Conc. (%)
PA	0.968	16.8	14.2	0.933	11.1	4.62
Cellulose	0.996	22.04	37.6	0.993	21.6	49.99
Lactose	0.995	25.41	48.16	0.995	25.1	44.88

Distribution des composés

Exploration des solutions

- Concentrations moyennes

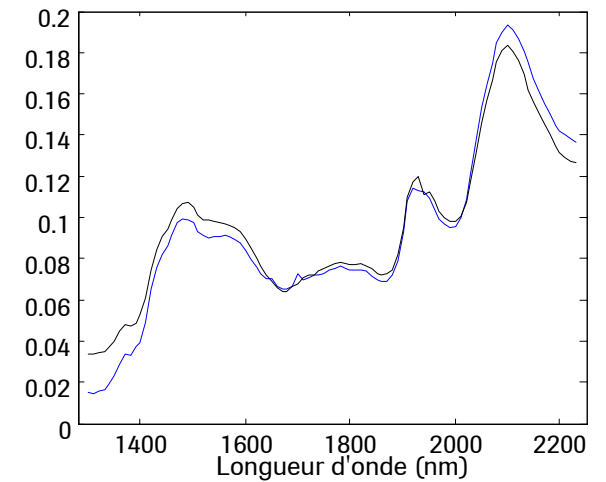
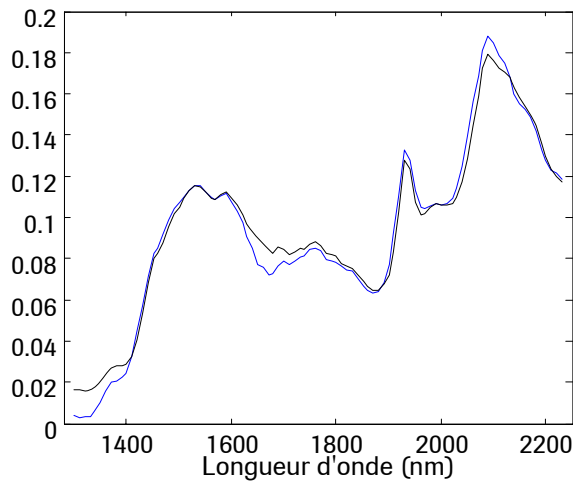
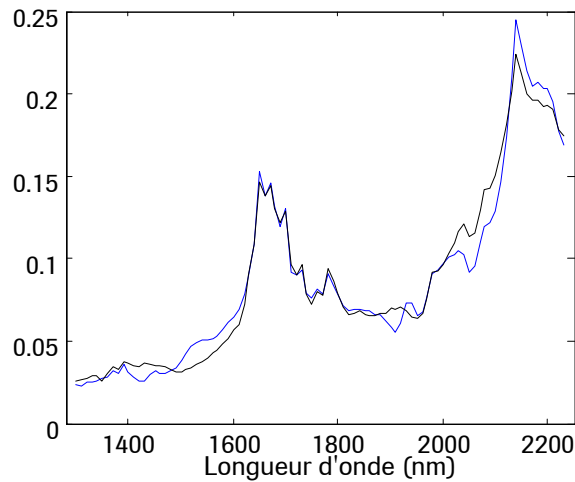


	Avant introduction des concentrations			Après introduction des concentrations		
	Corrélation	RSD	Conc. (%)	Corrélation	RSD	Conc. (%)
PA	0.968	16.8	14.2	0.933	11.1	4.62
Cellulose	0.996	22.04	37.6	0.993	21.6	49.99
Lactose	0.995	25.41	48.16	0.995	25.1	44.88

Distribution des composés

Exploration des solutions

- Concentrations moyennes



	Avant introduction des concentrations			Après introduction des concentrations		
	Corrélation	RSD	Conc. (%)	Corrélation	RSD	Conc. (%)
PA	0.968	16.8	14.2	0.987	21.3	7.98
Cellulose	0.996	22.04	37.6	0.994	21.82	47
Lactose	0.995	25.41	48.16	0.994	24.98	45.17

Distribution des composés

Bilan

- L'imagerie proche infrarouge permet la détermination de la distribution des composés chimiques
- Si les spectres de références sont connus l'algorithme PLS-DA donne une bonne estimation des cartes de distributions
- Sans les spectres de références, les algorithmes MCR-ALS et PMF peuvent être utilisés
- Les matrices de rotation et l'utilisation de la concentration moyenne des composés améliorent l'extraction

Introduction

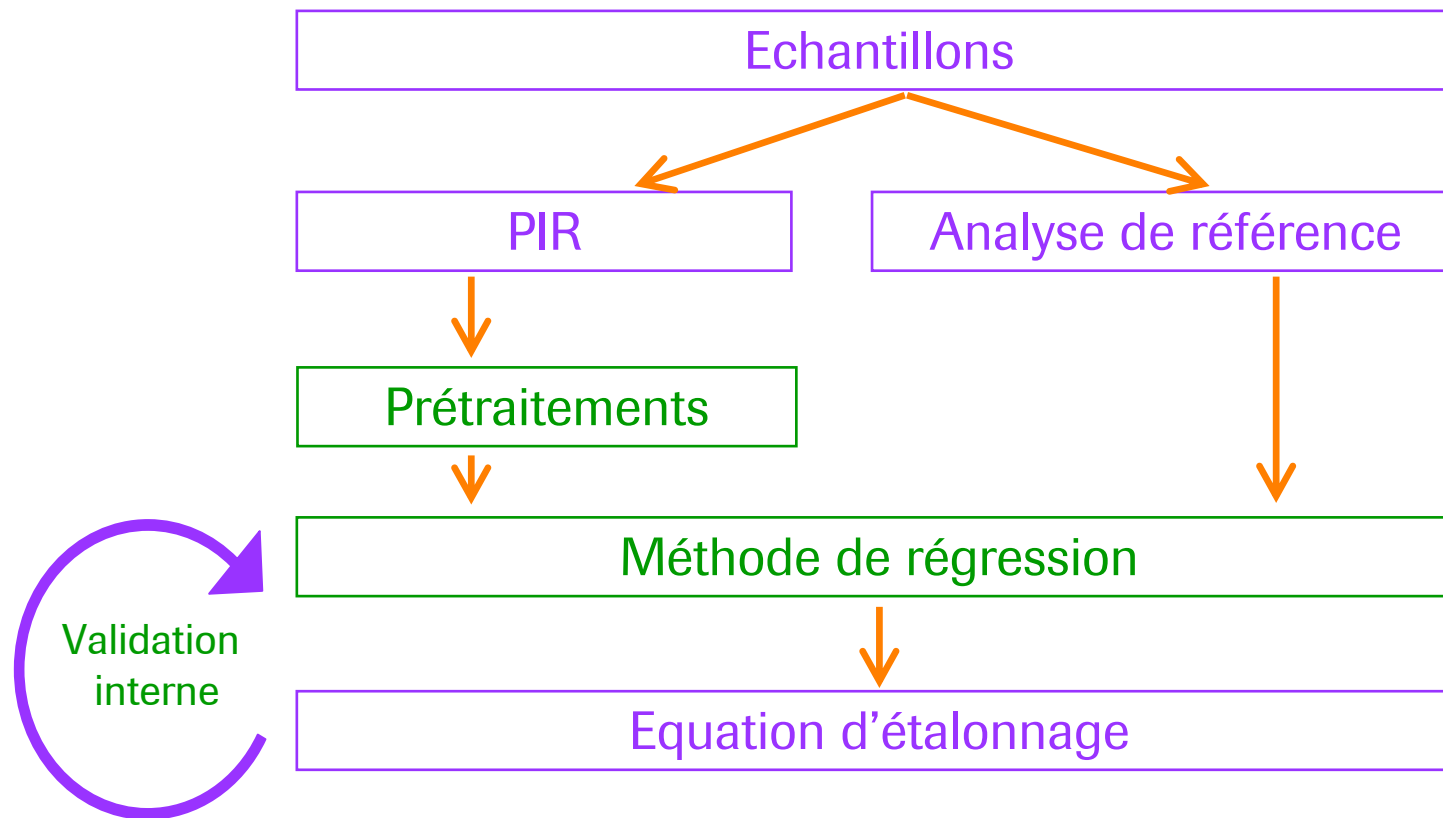
Distribution des composés

Détermination de la teneur en principe actif

Conclusions

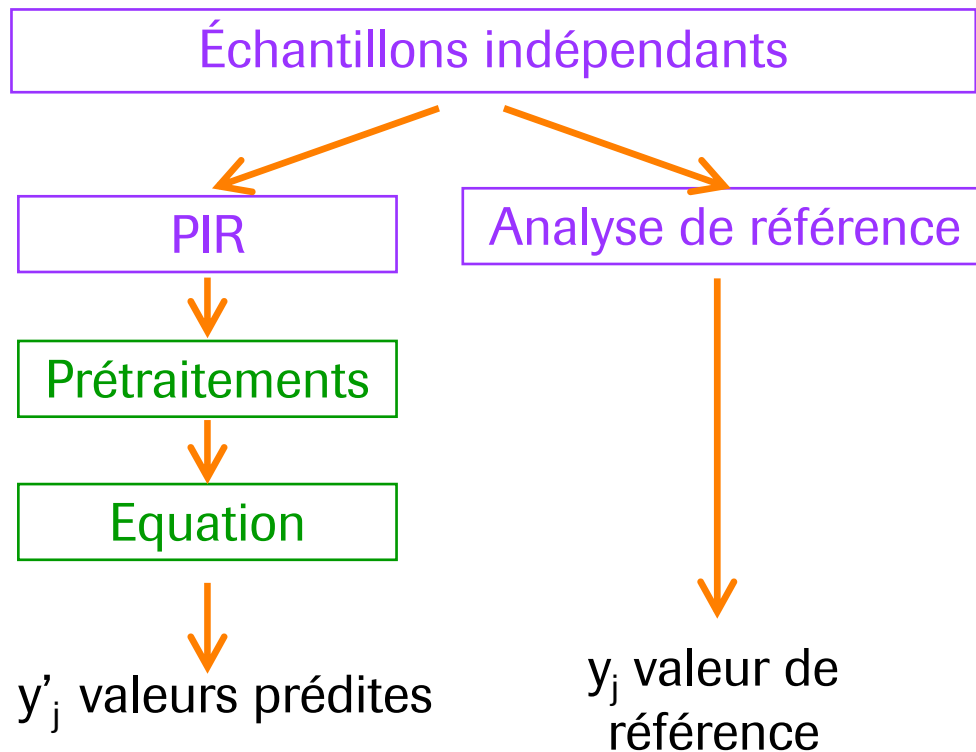
Principe de l'analyse quantitative

Première étape : étalonnage



Principe de l'analyse quantitative

Seconde étape : validation



Evaluation du modèle.

$$SEP = \sqrt{\frac{\sum_{i=1}^n (y_i - y'_i)^2}{n}}$$

n = nombre d'échantillons

Linéarité de la droite d'étalonnage:

- R²
- Pente
- Ordonnée à l'origine

Etude préliminaire : lots binaires

Sans a priori : algorithmes et échantillons

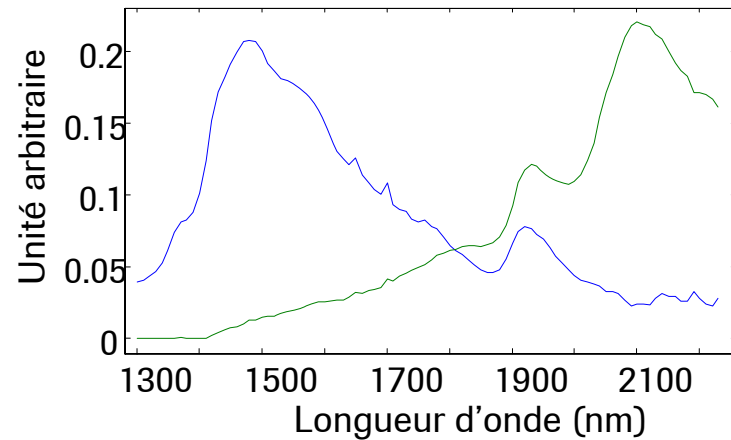
- Echantillons
 - Cellulose – Principe Actif : particules micronisées
 - Les comprimés sont produits dans le laboratoire
 - 0 à 100 % d'actif par pas de 10%
 - 44 comprimés sont analysés (4 gammes différentes)
 - Référence : pourcentage massique des mélanges
- Algorithmes
 - Algorithmes de résolution de courbes multivariées (MCR-ALS / PMF)
 - Segmentation : K-moyennes avec deux classes

Etude préliminaire : lots binaires

Sans a priori : résultats

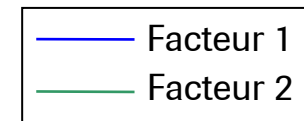
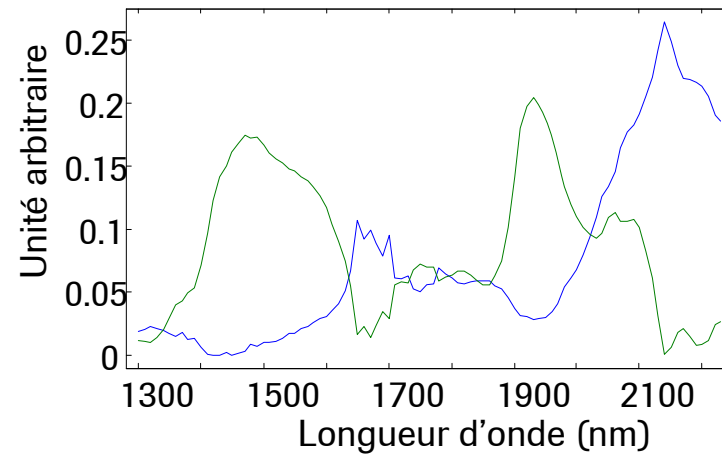
Facteurs extraits par MCR-ALS

20%

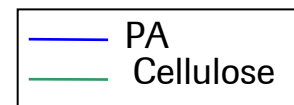
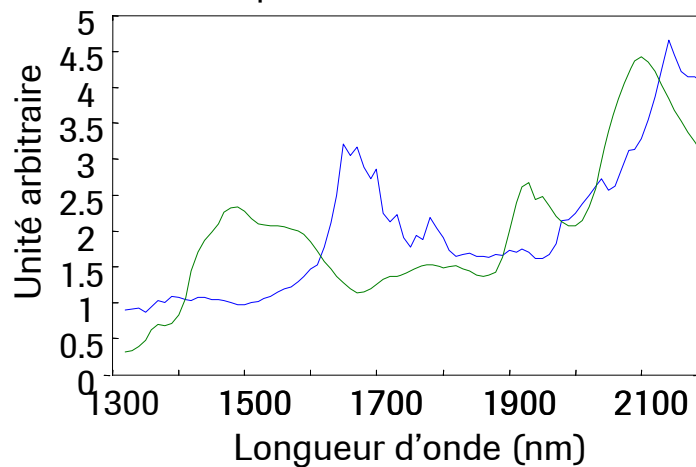


Facteurs extraits par MCR-ALS

30%



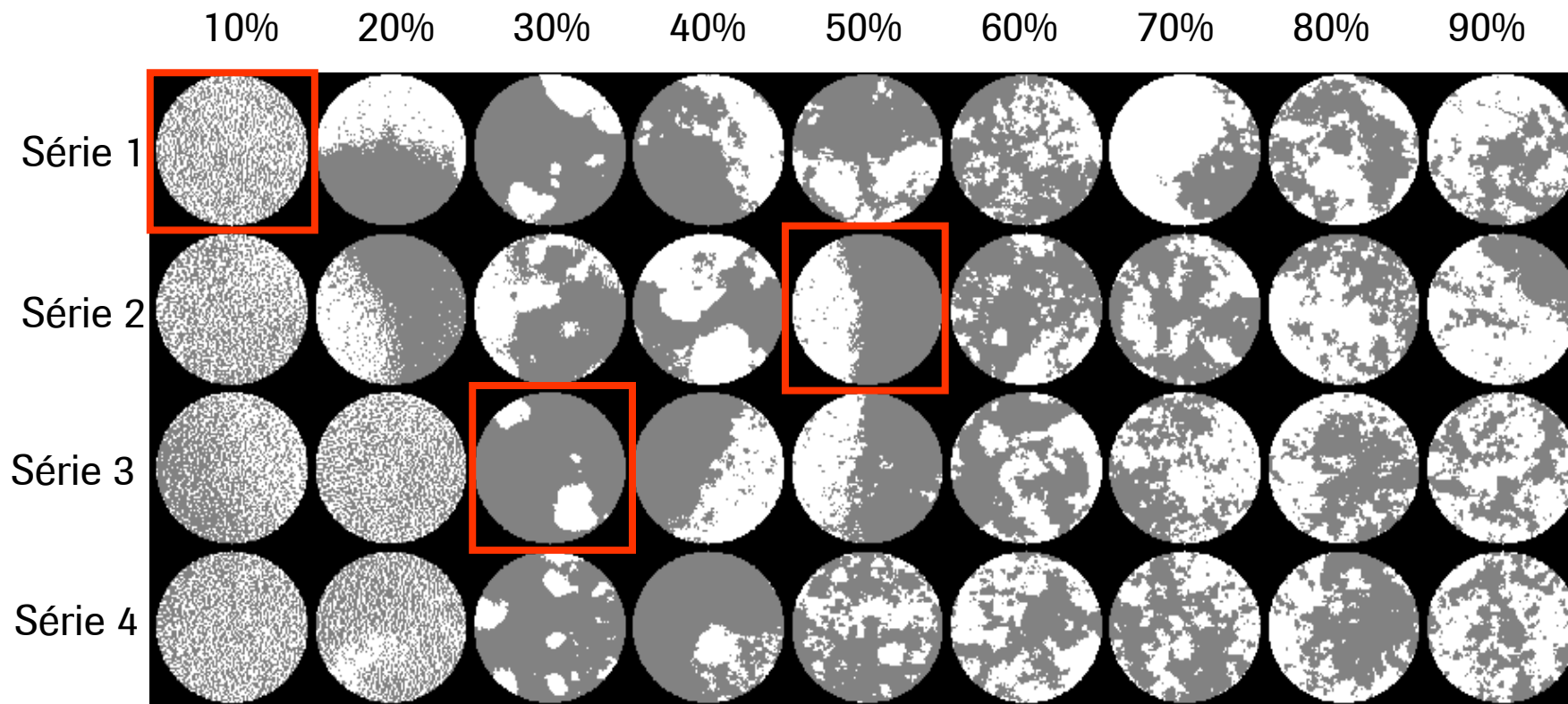
Spectres de référence



➔ Ambigüités rotationnelles empêchent une bonne extraction

Etude préliminaire : lots binaires

Sans a priori : résultats



Résolution spatiale effective n'est pas suffisamment fine pour une segmentation correcte des particules micronisées

Formulation réelle

- La quantification sans a priori est difficile pour des particules micronisées
- Pour quantifier l'actif présent à faible dosage et micronisé : introduction des informations supplémentaires.
- Formulation pharmaceutique :
 - Principe actif, cellulose, lactose, magnésium stéarate, talc et deux colorants
 - Concentration en principe actif varie de 0% à 10% par pas de 1%
 - Les comprimés sont produits dans le laboratoire
 - 44 comprimés sont analysés (4 gammes différentes)

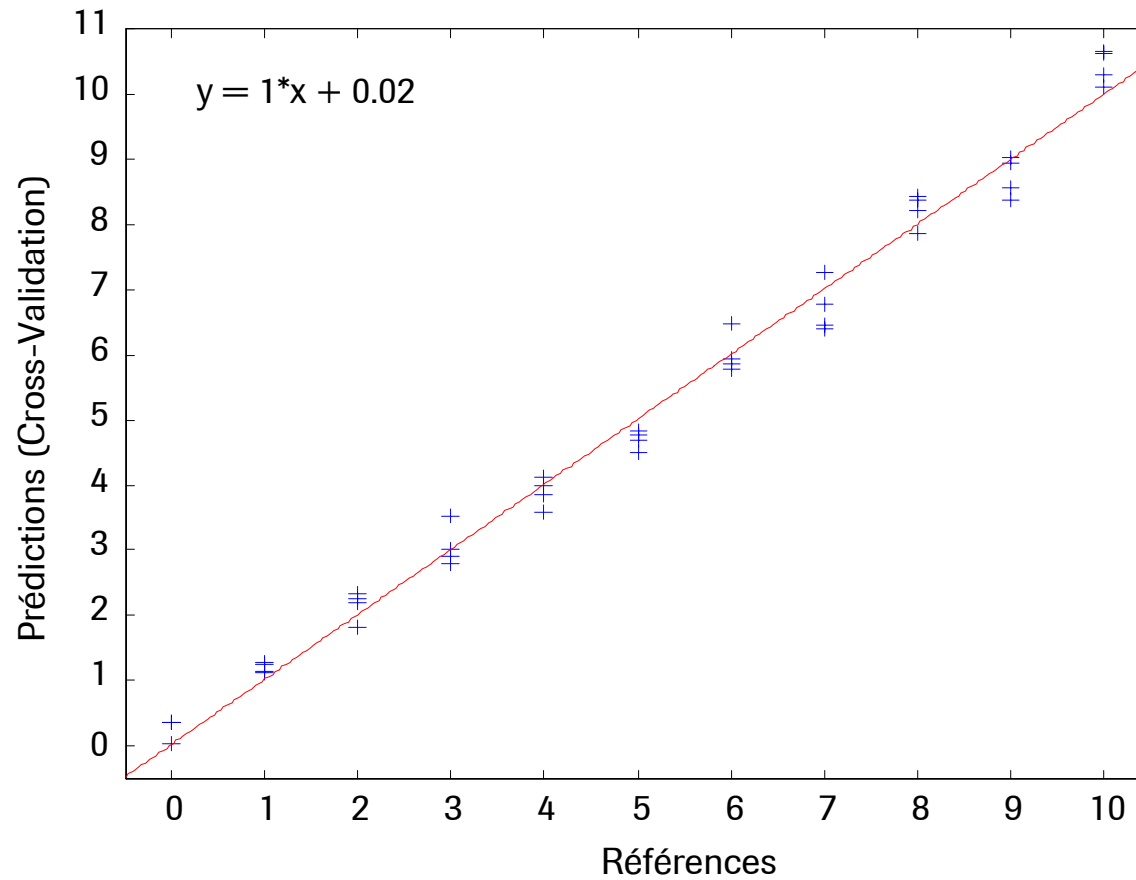
Formulation réelle

Avec a priori : gamme de concentration

- Algorithme des moindres carrés partiels (PLS)
- Valeurs de référence : pourcentage massique des échantillons
- Le modèle est construit avec les spectres moyens de chaque comprimé
- La teneur en actif de chaque pixel est ensuite prédite pour former les cartes de distributions

Formulation réelle

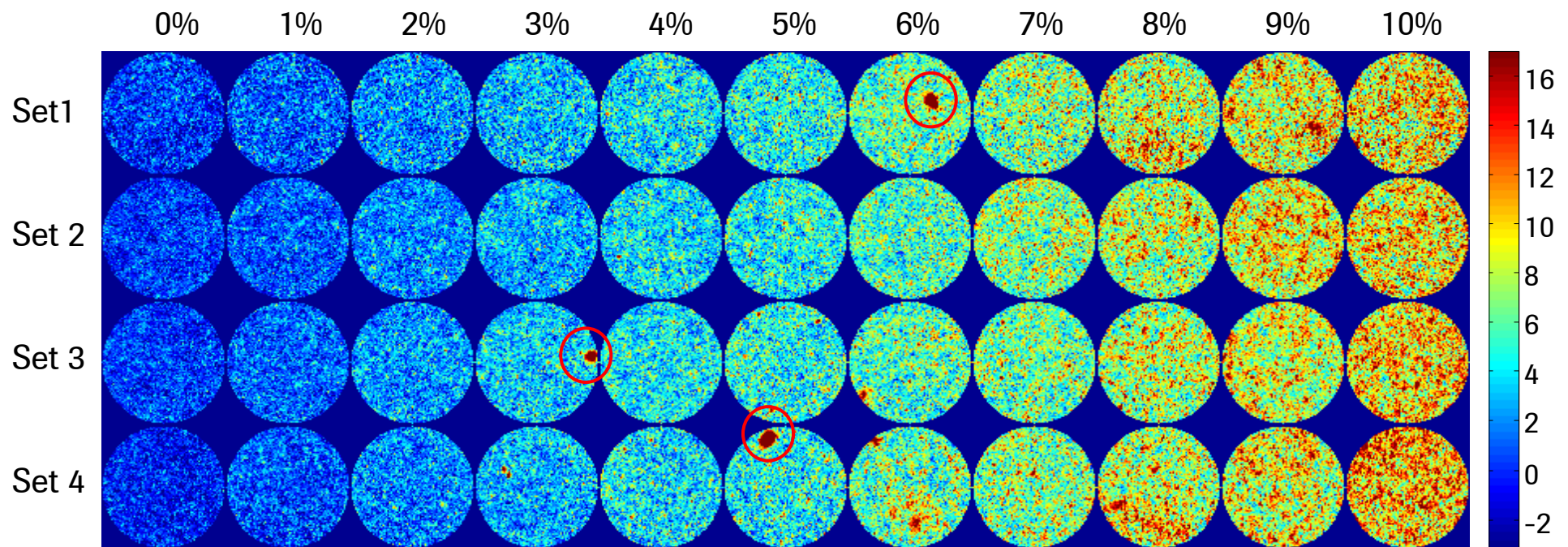
Avec a priori : gamme de concentration



SECV : 0.34%

Formulation réelle

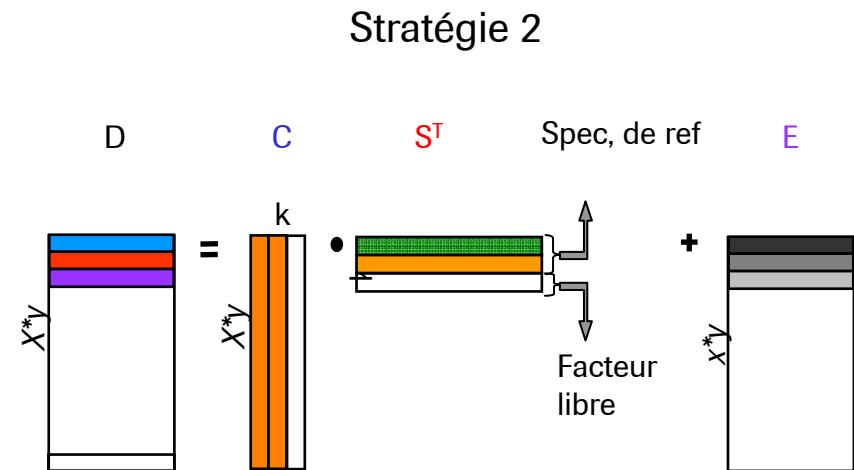
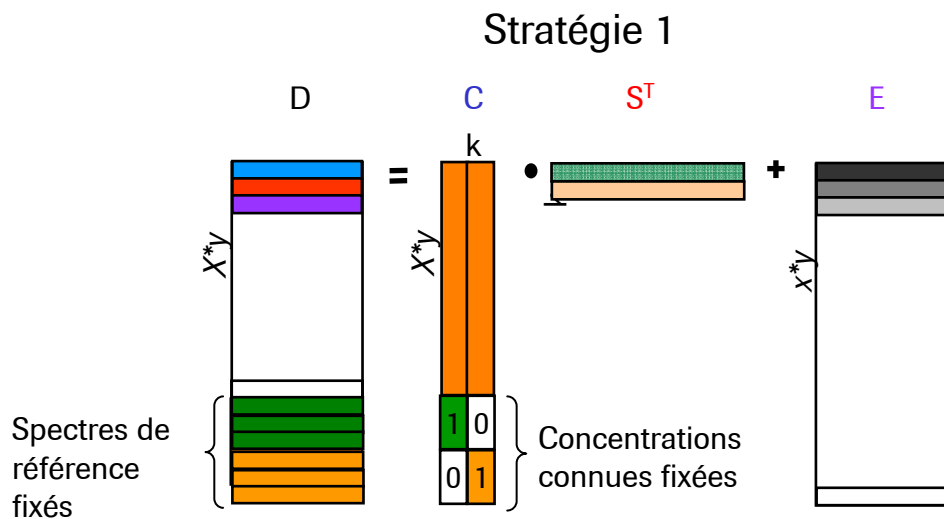
Avec a priori : gamme de concentration



Formulation réelle

Avec a priori : spectres de références

- PLS-DA, moindres carrés
- MCR-ALS augmenté



Formulation réelle

Avec a priori : spectres de références

Prétraitements : normalisation SNV et dérivée seconde

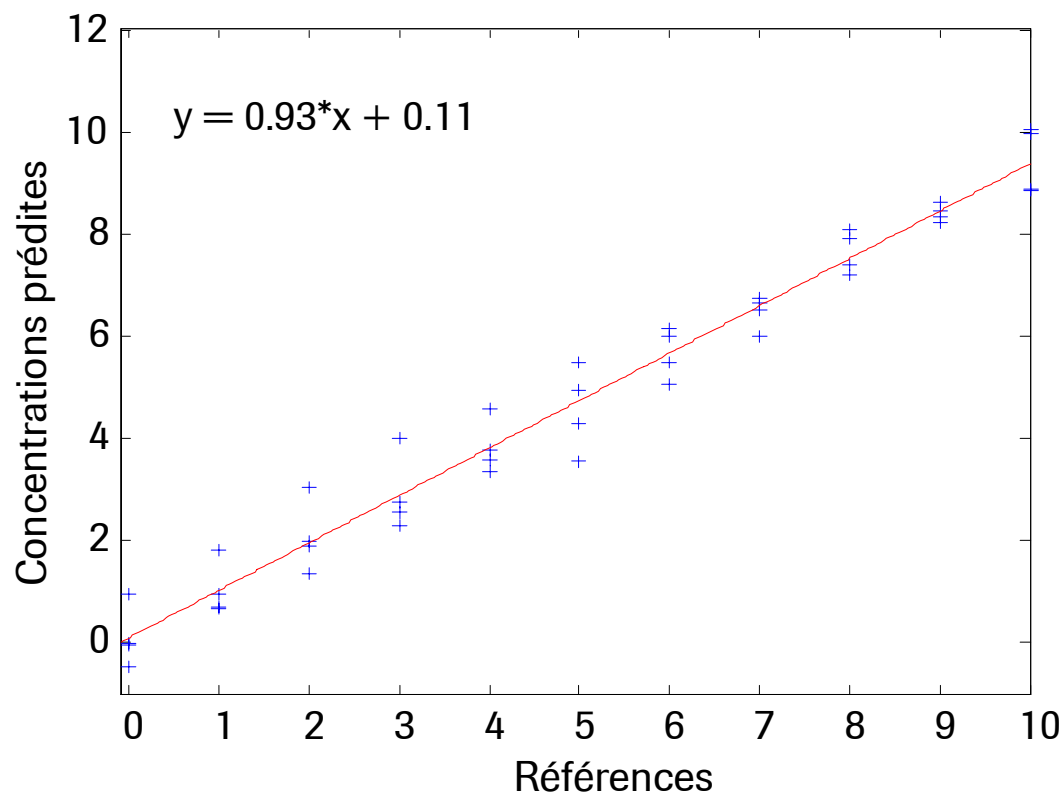
	PLS-DA	CLS	MCR-ALS (1)	MCR-ALS (2)
Pente	0.926	1.009	0.920	0.899
Ordonnée à l'origine	0.11	-1.128	0.698	0.768
R ²	0.971	0.967	0.959	0.939
SEP (%)	0.62	1.235	0.719	0.835

➔ La meilleure méthode est la PLS-DA

Formulation réelle

Avec a priori : spectres de références

PLS-DA, dérivée seconde, 5 variables latentes, 3 références



SEP : 0.62%

Formulation réelle

Bilan

- La quantification d'un Principe Actif sans aucun a priori est difficile
- La quantification précise avec une gamme de concentration est possible.
- Avec les spectres de références : algorithme PLS-DA
- L'imagerie permet l'extraction simultanée de la concentration et des cartes de distributions

Introduction

Distribution des composés

Détermination de la teneur en principe actif

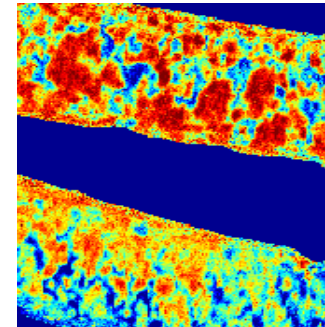
Conclusions

Conclusions

- Extraction de la distribution des composés :
 - Avec références connues : analyse à une longueur d'onde, PLS-DA
 - Sans a priori : MCR-ALS et PMF avec les outils d'exploration des solutions (matrices de rotations, concentrations moyennes)
- Quantification :
 - Avec gamme de concentration : PLS
 - Avec spectres de références : PLS-DA

Conclusions

- L'imagerie proche infrarouge est un outil innovant ¹:
 - Distribution des composés ²
 - Compréhension du procédé ³
 - Quantification de l'actif ⁴



- C'est un outil *Process Analytical Technology*

1 C. Gendrin, Y. Roggo, C. Collet, Vibrational chemical imaging and chemometrics for pharmaceutical applications: A review, Journal of Pharmaceutical and Biomedical Analysis, Vol 48 (2008), p. 533-553

2 C. Gendrin, Y. Roggo, C. Collet, Self modelling curve resolution of Near Infrared Imaging Data, Proceedings of the ICNIRS 2007, Umea, Sweden, Journal of Near Infrared Spectroscopy , Vol 16 (2008), Issue 3, p. 151-157

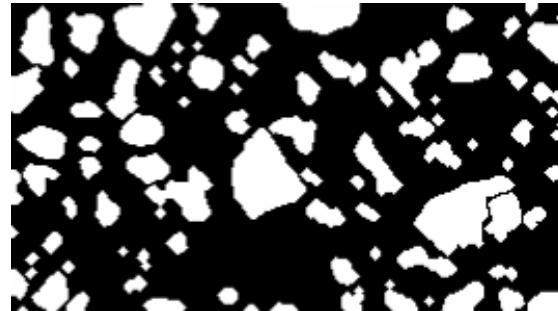
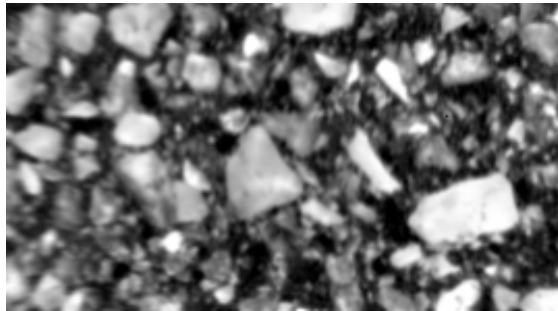
3 C. Gendrin, Y. Roggo, C. Spiegel, C. Collet, Monitoring Galenical Process Development by NIR Chemical Imaging: one case study, European Journal of Pharmaceutics and Biopharmaceutics, Vol 68 (2008), Issue 3, p. 376-385

4 C. Gendrin, Y. Roggo, C. Collet, Content uniformity of pharmaceutical solid dosage forms by near infrared hyperspectral imaging: A feasibility study, Talanta, Vol 73 (2007), p. 733-741

Perspectives

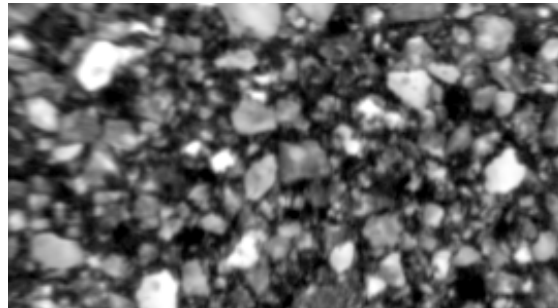
Applications en ligne : taille de particules

Lot 1



0.28 mm²

Lot 2



0.24 mm²

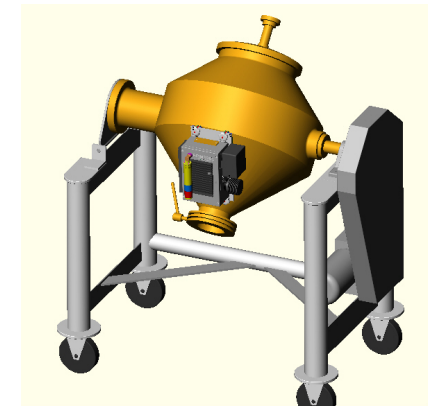
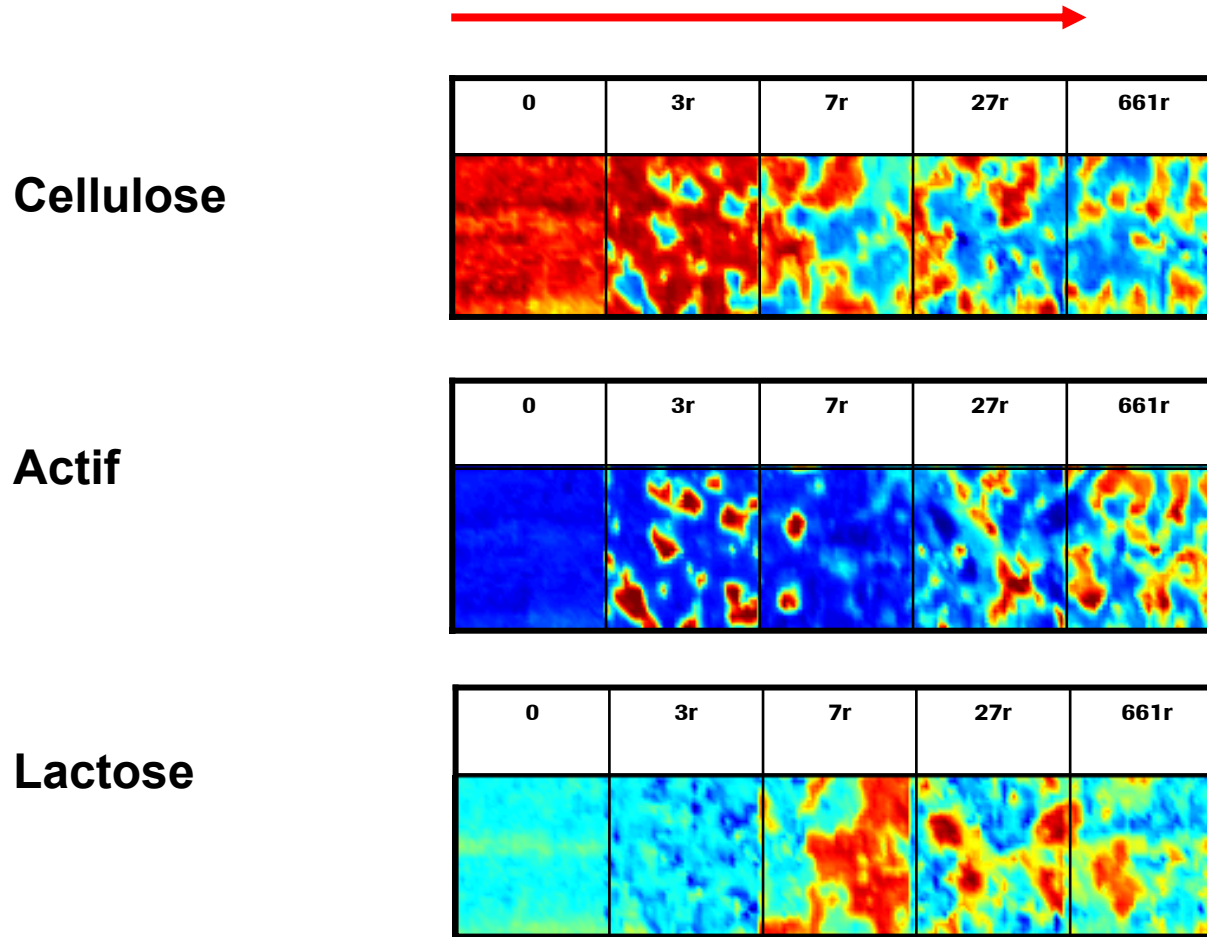
Image prétraitée

Image segmentée

Taille moyenne

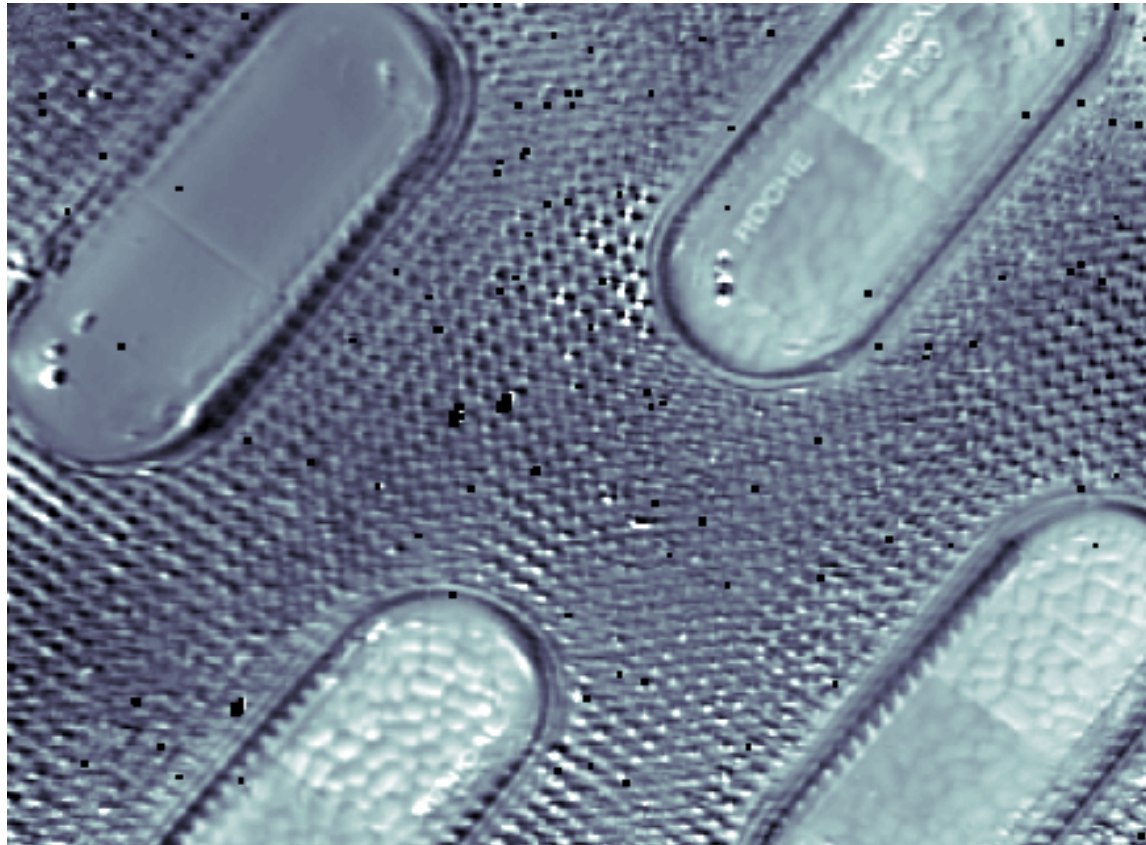
Perspectives

Applications en ligne : uniformité des mélanges



Perspectives

Applications en ligne : détection de capsules vides par imagerie



Merci de votre attention!!



Nous innovons la santé