



Université Claude Bernard  Lyon 1



INSTITUT DES
SCIENCES
ANALYTIQUES



Les apports du numérique dans l'analyse physicochimique de produits alimentaires : la chimiométrie.

De la stratégie expérimentale à l'analyse de données :
intelligence naturelle , intelligence artificielle ?



Yohann Clément & Pierre Lantéri



Co-funded by the Horizon 2020 Framework Programme of the European Union
under the grant N° 952306



Université Claude Bernard



Lyon 1



INSTITUT DES
SCIENCES
ANALYTIQUES



Le Rôle de la Chimiométrie dans les Sciences Analytiques

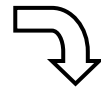
- Participer à la stratégie expérimentale pour structurer l'expérimentation (PEX)
- Participer à l'acquisition des données souvent complexes issues d'une ou plusieurs techniques analytiques
- Traiter les données brutes pour optimiser la matrice des données
- Traitement de l'information / résultats/ enrichissement de bases de données

Origine des données

Données Issues d'essais structurés a priori?



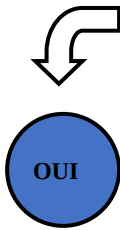
Plans d'Expériences



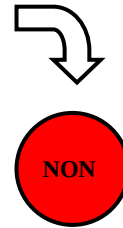
Analyse de Données

Analyse de données

Variables explicatives
et à expliquer séparées :



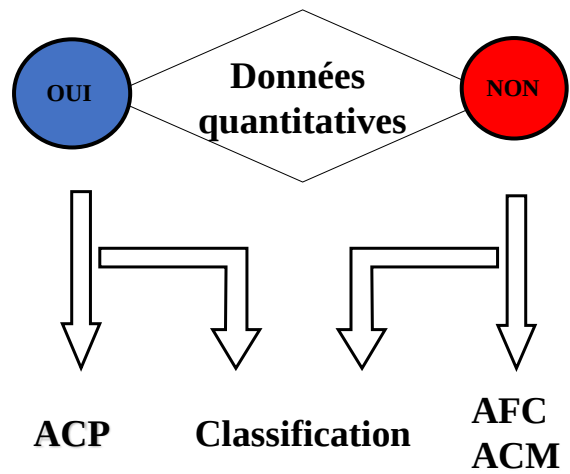
Méthodes explicatives



Méthodes descriptives

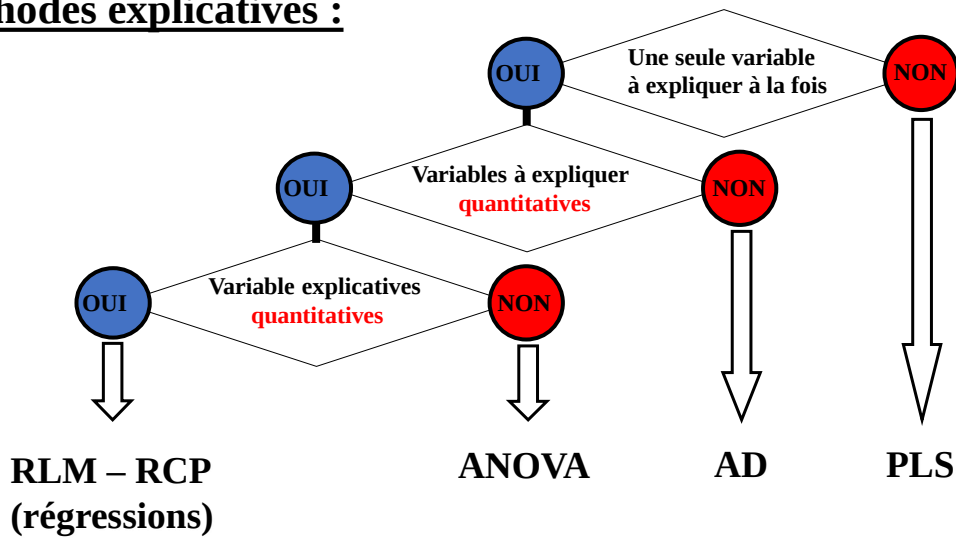
Analyse de données

Méthodes Descriptives :



Analyse de données

Méthodes explicatives :



Thème	Domaines	Méthodes/Outils
Analyses de Données	Analyses NIR, IR-FT, UV, Propriétés Physicochimiques Chromatographies HPLC,GC Spectrométrie Masse, RMN Environnement, Santé, Matériaux,...	- Analyses Statistiques, - Analyse Factorielle - Analyse en composantes principales (PCA) -Régressions RLM/PCR/PLS, -Réseaux neuronaux -QSAR/QSPR : Relation Structures- Activité,Propriétés,
Calibration/ Etalonnage	Spectroscopie : UV, IR-FT, NIR Chromatographies Analyses chimiques ...	-Régressions MLR, PCR, PLS -Réseaux neuronaux -PCA
Classification	Identification Adultération Contrôle qualité Métabolomique ...	- Réseaux neuronaux - PCA - Analyse Discriminante - K means ...
« intelligence artificielle »	Analyses Chimiques, pH métrie, QSAR, Etalonnage, Spectrométrie, PLS/calibration ...	-Réseaux neuronaux -Système flou -Algorithme Génétique

Les données: de l'acquisition au traitement

Stratégies
Expérimentales

Criblage
Hadamard
Factorielle
Fractionnaire
...

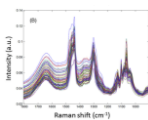
Surface de réponse
Plan composite
Réseau de Dohert
Box Behnken
...

Mélange
Simplex
Réseau de Scheffé
...

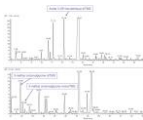
Gestion de contraintes
Algorithmes d'échanges
ASCA
Space Filling Design
...

Acquisitions
analytiques

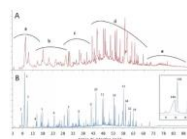
Spectroscopie
IR - RAMAN



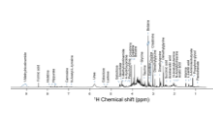
Spectrométrie
de masse



Chromato.



Spectroscopie
RMN



**Spectroscopie
vibrationnelle**

Spectrométrie de masse

**Chromatographies,
couplages**

RMN

Traitement des données

OUTILS CHIMIOMETRIQUES

Mise en forme

Normalisation

...

Filtrage

Alignement de variables
Correction de ligne de
base. ...

Transformation

Log10
1/T. ...

Changement d'échelles

Centrage
Standardisation ...

Statistiques

Univariate Tests

t-student
Correlation
...

Description

PCA
ICA
...

Modelisation

PLS
PLS-DA
OPLS
...

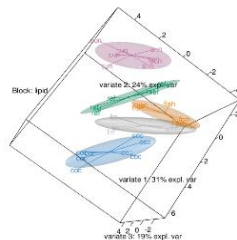
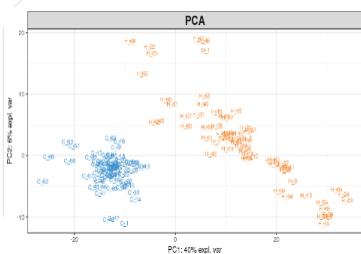
Multi-blocks method

AFM
PARAFAC
COMDIM
...

Machine Learning

SVM
Random Forest
Réseaux de neurones
...

Représentations



Analyse en composante Principale ACP

Analyse en Composantes Principales (ACP) :

A partir d'un jeu de données contenant un très grand nombre de variables, on désire connaître lesquelles sont les plus importantes, et comment réduire ce jeu de données afin de le représenter de manière simple sur 2 ou 3 axes.

L'ACP est une méthode de réduction de dimension qui va permettre de transformer des variables très corrélées en nouvelles variables décorrélées les unes des autres.

Il s'agit de résumer l'information qui est contenue dans une base de données en un certain nombre de variables synthétiques appelées : **Composantes principales**. L'idée est ensuite de pouvoir projeter ces données sur l'hyperplan le plus proche afin d'avoir une représentation simple de nos données.

Qui dit réduction de dimension dit perte d'informations. C'est là tout l'enjeu que représente une Analyse en Composantes principales. Il faut pouvoir réduire la dimension de nos données tout en conservant un maximum d'informations.

Analyse en composante Principale ACP

L'ACP est une méthode **exploratoire / descriptive**

- Exploration / description de données multivariées avec de grandes dimensions.
- Elle permet la détermination des liens entre variables, l'identification des similitudes entre individus ou la mise en évidence de valeurs aberrantes ...
- Elle permet la préparation / réduction ou nettoyage des données avant une nouvelle analyse chimiométrique.
- Elle peut préparer à une future modélisation comme la Régression sur Composantes Principales (PCR)

Analyse en Composante Principales : résumé de la méthode

Collecte des données



Prétraitement des données : Normalisation, Centrage ...



Calculs de l'ACP:

Représentation matricielle

Calculs des valeurs propres et vecteurs propres de la matrice de corrélation



Sélection d'un nombre réduit de Composantes Principales (PC)

Représentations graphiques dans un espace de quelques dimensions (2,3...)

Projections des individus (scores) et des variables (loadings)

Analyse de miels
à l'aide la chimiométrie
Analyse en Composante Principale
(ACP)

13

Analyse de miels monofloraux

Analyse des miels : détermination de l'origine florale?

**Amino
acides**

**Acide
Gluconique**

**Sucres :
Fructose,
Glucose, ...**

Variétés de miel

Acacia
Châtaignier
Lavande
Sapin
Tournesol



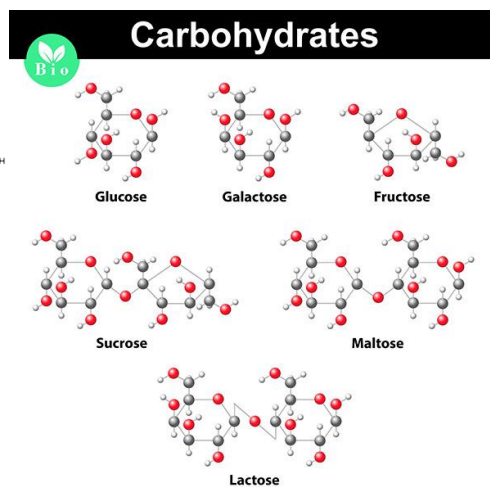
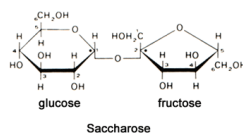
Analyse de miels monofloraux

Les descripteurs mesurés et dosés sont :

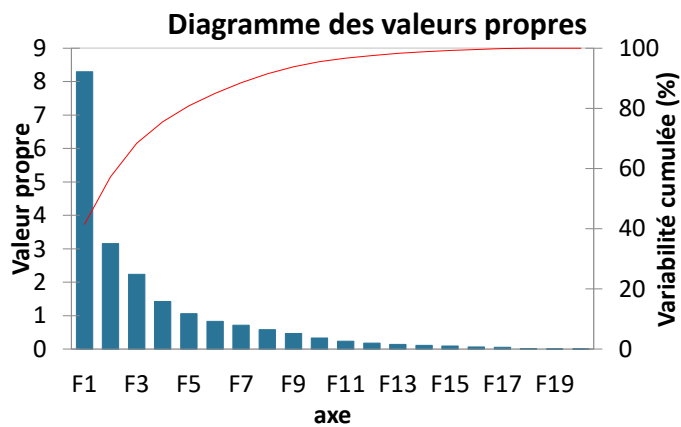
- Les sucres
- Les acides aminés

Analyse de miels monofloraux – ACP- sucres

Variables
SUCRES
QFructose (g/kg)
QGlucose (g/kg)
Fructose
Saccharose
Maltose
Maltulose
Turanose
Tréhalose
Laminaribiose
Mélibiose
Isomaltose
Gentiobiose
Raffinose
Néo-kestose
1-Kestose



Analyse de miels monofloraux – ACP -Sucres

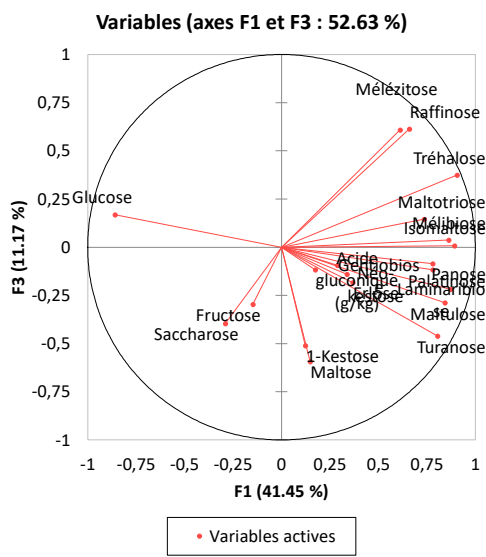
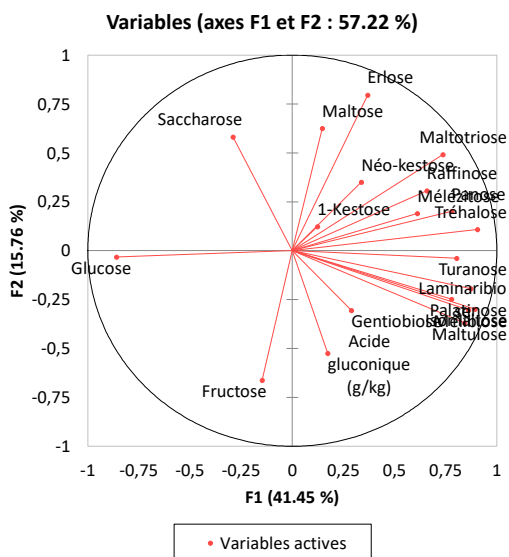


	F1	F2	F3	F4	F5
Valeur propre	8,2	3,1	2,2	1,4	1,0
Variabilité (%)	41,4	15,7	11,1	7,1	5,2
% cumulé	41,4	57,2	68,3	75,4	80,7

17

Analyse de miels monofloraux – ACP - Sucres

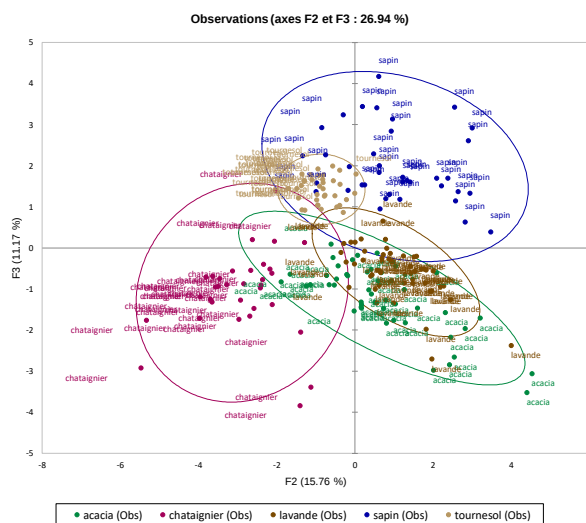
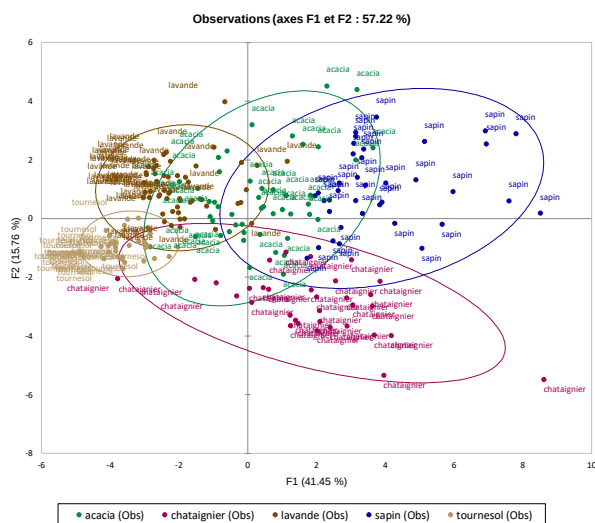
Projection des variables (loadings)



18

Analyse de miels monofloraux – ACP - Sucres

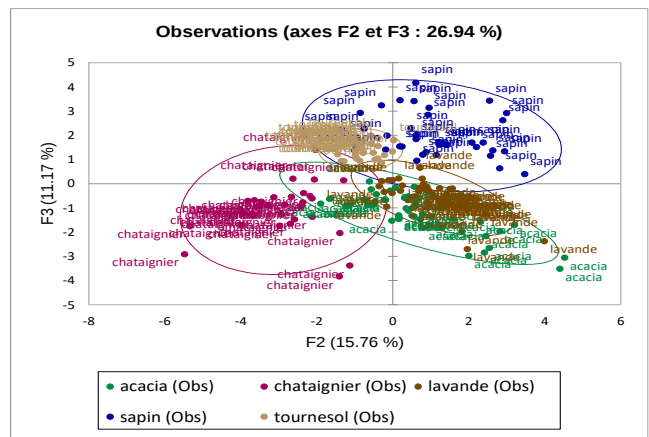
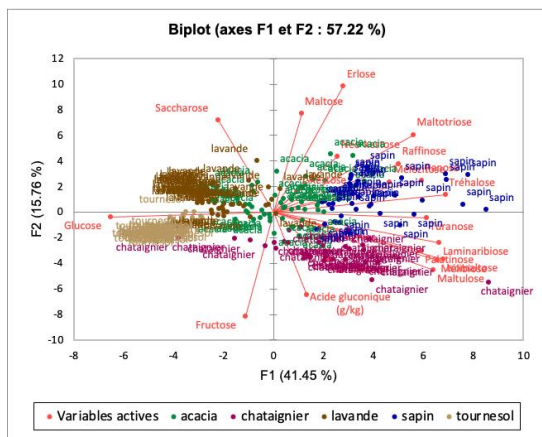
Projection des individus (scores) :



19

Analyse de miels monofloraux – ACP -Sucres

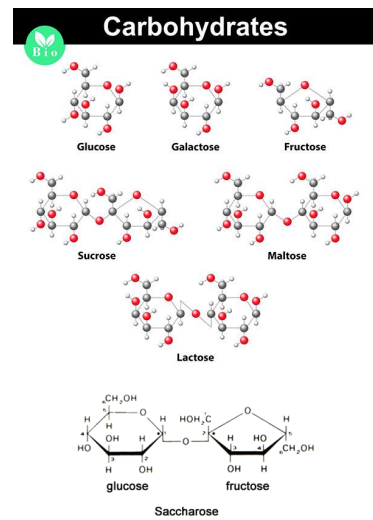
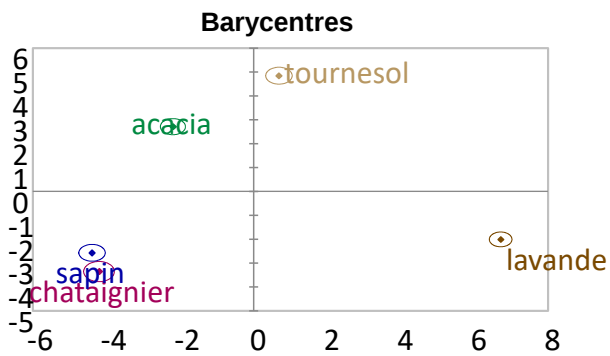
Biplot – projection sur le même graphique des variables et des individus (scores et des loadings):



Analyse des miels monofloraux: ANALYSE DISCRIMINANTE (AD) via les sucres

21

Analyse de miels monofloraux – AD - Sucres



22

Analyse de miels monofloraux – AD - Sucres

Jeu de données divisée en deux: Un jeu d'apprentissage et un jeu de validation externe

Matrice de confusion pour l'échantillon d'apprentissage :							
de \ à	acacia	chataignier	lavande	sapin	tournesol	Total	% correct
acacia	35	0	0	0	1	36	97,22%
chataignier	0	22	0	0	1	23	95,65%
lavande	0	0	41	0	1	42	97,62%
sapin	0	0	0	30	0	30	100,00%
tournesol	0	0	0	0	25	25	100,00%
Total	35	22	41	30	28	156	98,08%

Matrice de confusion pour l'échantillon de validation :							
de \ à	acacia	chataignier	lavande	sapin	tournesol	Total	% correct
acacia	11	0	0	0	0	11	100,00%
chataignier	0	10	0	0	0	10	100,00%
lavande	0	0	10	0	1	11	90,91%
sapin	0	0	0	7	0	7	100,00%
tournesol	0	0	0	0	11	11	100,00%
Total	11	10	10	7	12	50	98,00%

23

Analyse de miels monofloraux – AD - Sucres

Jeu de données monobloc: validation par validation croisée

Matrice de confusion pour les résultats de la validation croisée :							
de \ à	acacia	chataignier	lavande	sapin	tournesol	Total	% correct
acacia	34	0	1	0	1	36	94,44%
chataignier	0	22	0	0	1	23	95,65%
lavande	0	0	41	0	1	42	97,62%
sapin	0	0	0	30	0	30	100,00%
tournesol	0	0	0	0	25	25	100,00%
Total	34	22	42	30	28	156	97,44%

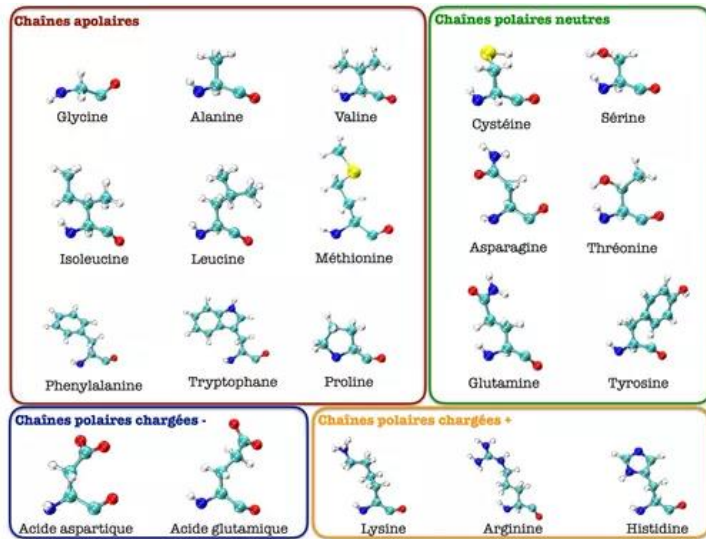
Analyse de miels monofloraux – Acides Aminés

Variables

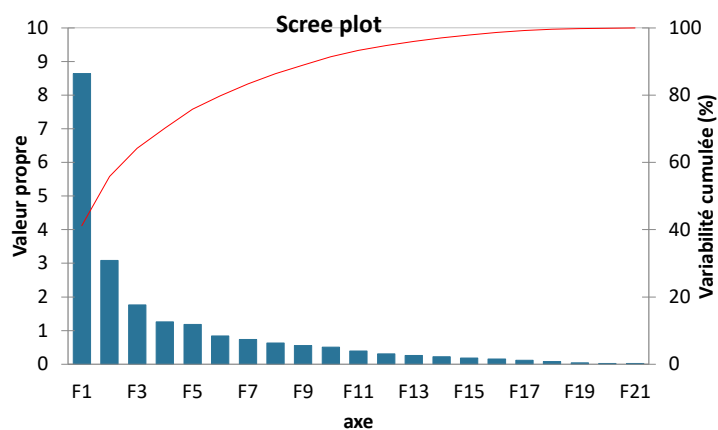
ACIDES AMINES

a aspartique
a glutamique
asparagine
sérine
glutamine
histidine
glycine
thrénine
alanine
arginine
valine
phénylalanine
isoleucine
leucine
lysine

LES ACIDES AMINES



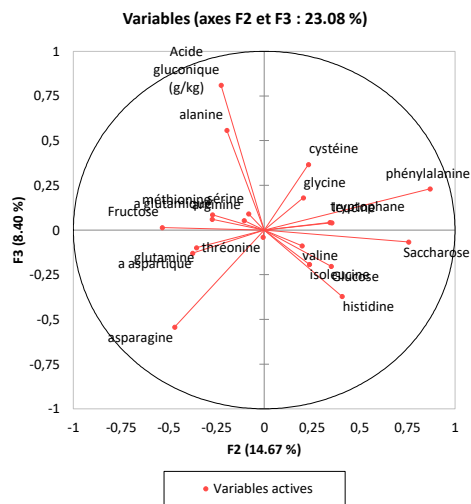
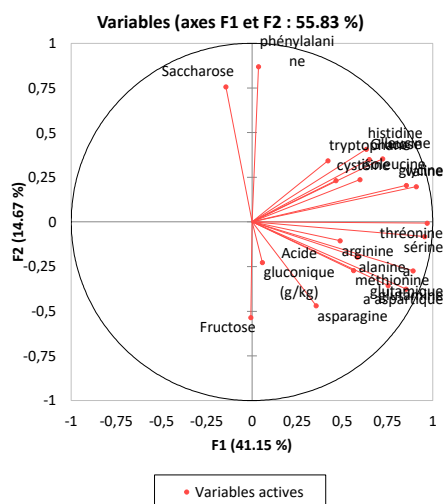
Analyse de miels monofloraux – ACP – Acides Aminés



	F1	F2	F3	F4	F5
Valeur propre	8,6	3,0	1,7	1,2	1,1
Variabilité (%)	41,1	14,6	8,4	5,9	5,6
% cumulé	41,1	55,8	64,2	70,2	75,8

Analyse de miels monofloraux – ACP- Acides Aminés

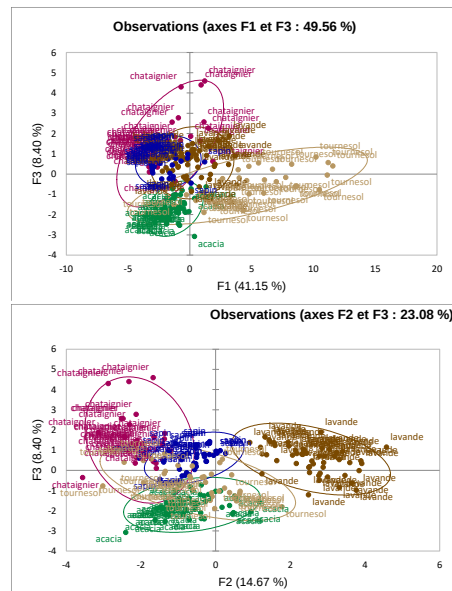
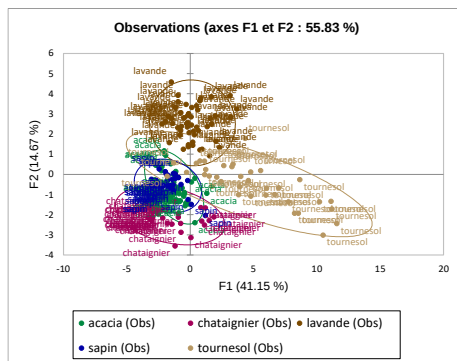
Projection des variables:



27

Analyse de miels monofloraux – AD – Acides Aminés

Projection des individus:

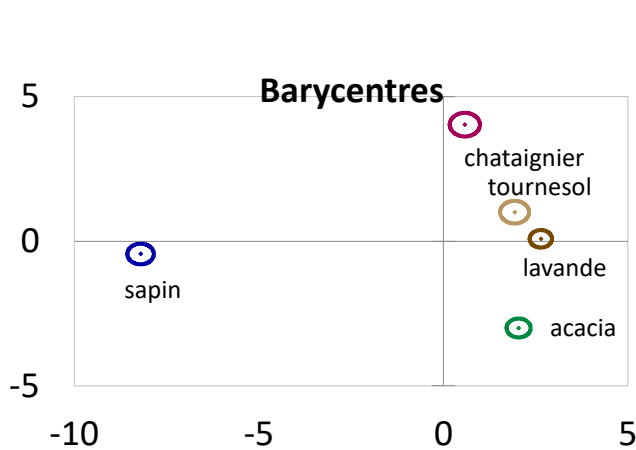


28

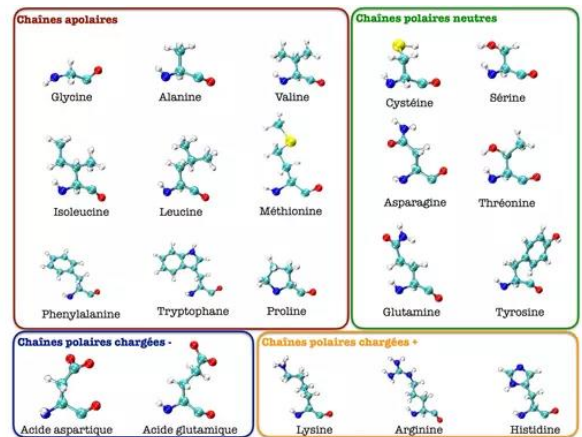
Analyse des miels monofloraux:
ANALYSE DISCRIMINANTE (AD)
via les acides aminés

29

Analyse de miels monofloraux – AD- Acides Aminés



LES ACIDES AMINES



Analyse de miels monofloraux – AD- Acides Aminés

Matrice de confusion pour l'échantillon d'apprentissage :							
de \ à	acacia	chataignier	lavande	sapin	tournesol	Total	% correct
acacia	38	0	0	0	0	38	100,00%
chataignier	0	26	0	0	0	26	100,00%
lavande	0	0	36	0	0	36	100,00%
sapin	0	0	0	28	0	28	100,00%
tournesol	0	0	0	0	28	28	100,00%
Total	38	26	36	28	28	156	100,00%

Matrice de confusion pour l'échantillon de validation :							
de \ à	acacia	chataignier	lavande	sapin	tournesol	Total	% correct
acacia	8	0	0	0	1	9	88,89%
chataignier	1	6	0	0	0	7	85,71%
lavande	0	0	17	0	0	17	100,00%
sapin	0	0	0	9	0	9	100,00%
tournesol	0	0	0	0	8	8	100,00%
Total	9	6	17	9	9	50	96,00%

31

Analyse de miels monofloraux – AD- Acides Aminés

Matrice de confusion pour les résultats de la validation croisée :							
de \ à	acacia	chataignier	lavande	sapin	tournesol	Total	% correct
acacia	38	0	0	0	0	38	100,00%
chataignier	0	26	0	0	0	26	100,00%
lavande	0	0	35	0	1	36	97,22%
sapin	0	0	1	27	0	28	96,43%
tournesol	0	0	0	0	28	28	100,00%
Total	38	26	36	27	29	156	98,72%

Méthodes analytiques pour le contrôle de l'authenticité

33



Université Claude Bernard



Développement de méthodes analytiques et de banques de données pour le contrôle de l'authenticité de métabolites secondaires actifs issus de plantes aromatiques et médicinales

Thèse présentée par **Aurélien CUCHET**

Encadrée par : Dr. Hervé CASABIANCA et M. Patrick JAME



INSTITUT DES
SCIENCES
ANALYTIQUES

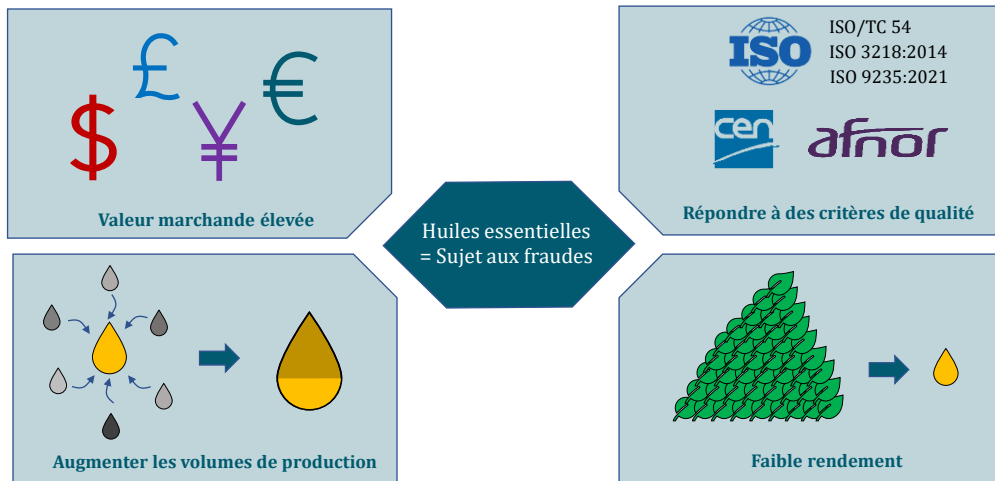
34

Présentation des extraits aromatiques naturels



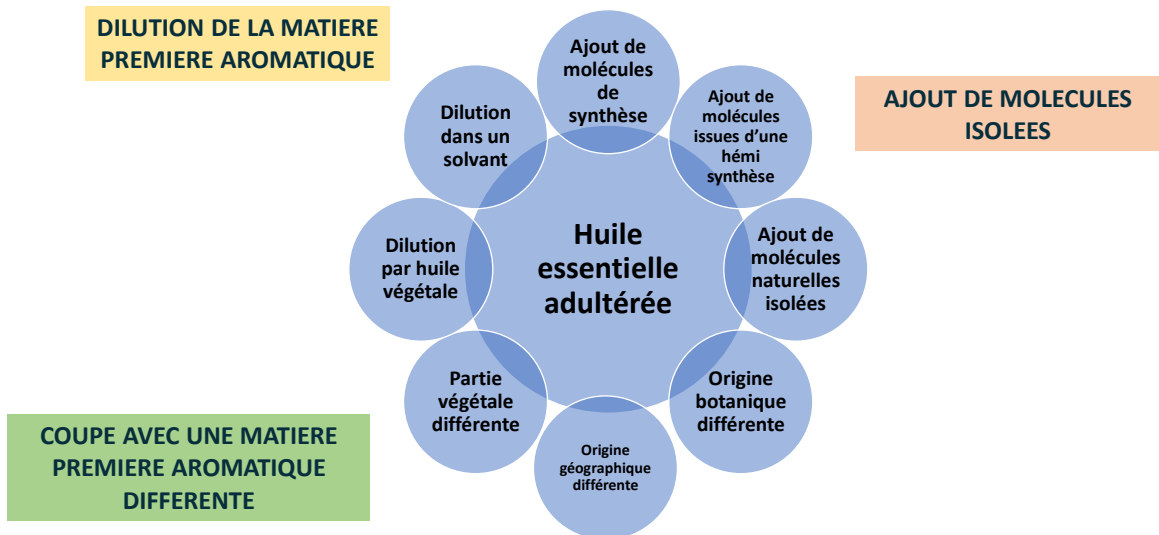
35

Les huiles essentielles sont la cible de nombreuses fraudes



36

Multiples adultérations potentielles



37

Contrôle des huiles essentielles

Techniques classiques

Exemples :

- Analyses par GC-FID et GC-MS
- Contrôles olfactifs
- Tests physicochimiques

Techniques avancées Combinaison de techniques

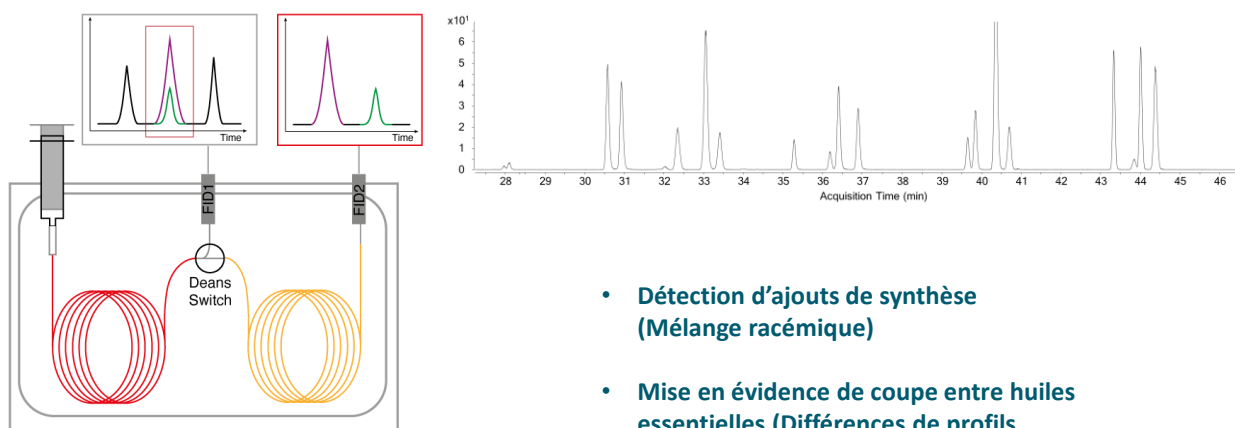
Exemples:

- Mise en évidence d'impuretés
- **Analyses énantiosélectives**
- **Analyses isotopiques**

38

ANALYSES ENANTIOSELECTIVES

Développement de méthodes GC multidimensionnelles

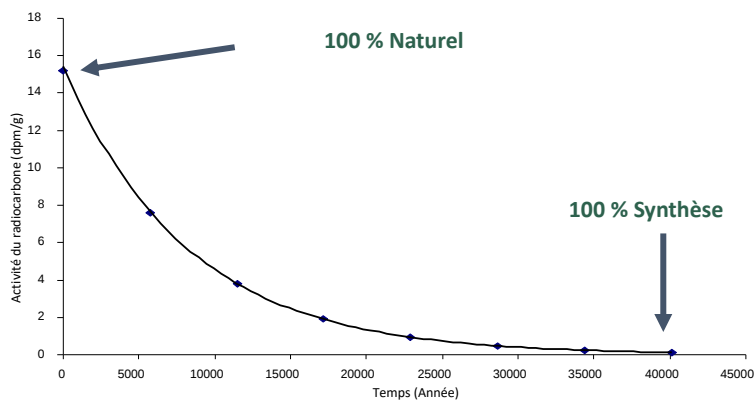


- **Détection d'ajouts de synthèse (Mélange racémique)**
- **Mise en évidence de coupe entre huiles essentielles (Différences de profils énantiomériques)**

39

ANALYSES ISOTOPIQUES

□ Analyse des isotopes radioactifs: détermination de l'activité du ^{14}C



L'authentification de la naturalité peut être réalisée en mesurant l'activité du ^{14}C

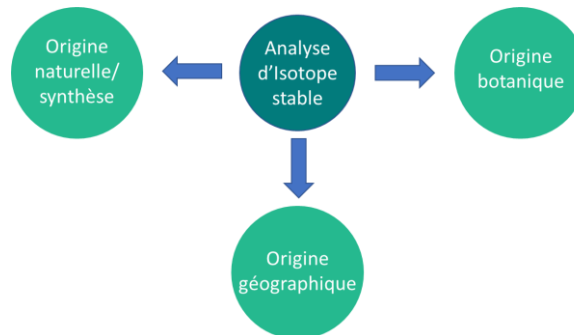
- 100% de carbone bio-sourcé -> Produit 100% naturel.
- 0% de carbone bio-sourcé -> Produit d'origine pétrochimique.
- Valeur comprise entre 0% et 100% -> Mélange d'un produit d'origine naturelle et d'origine synthétique.

L'activité radioactive est mesurée en désintégration par minute et par gramme de carbone (dpm/g).

40

ANALYSES ISOTOPIQUES

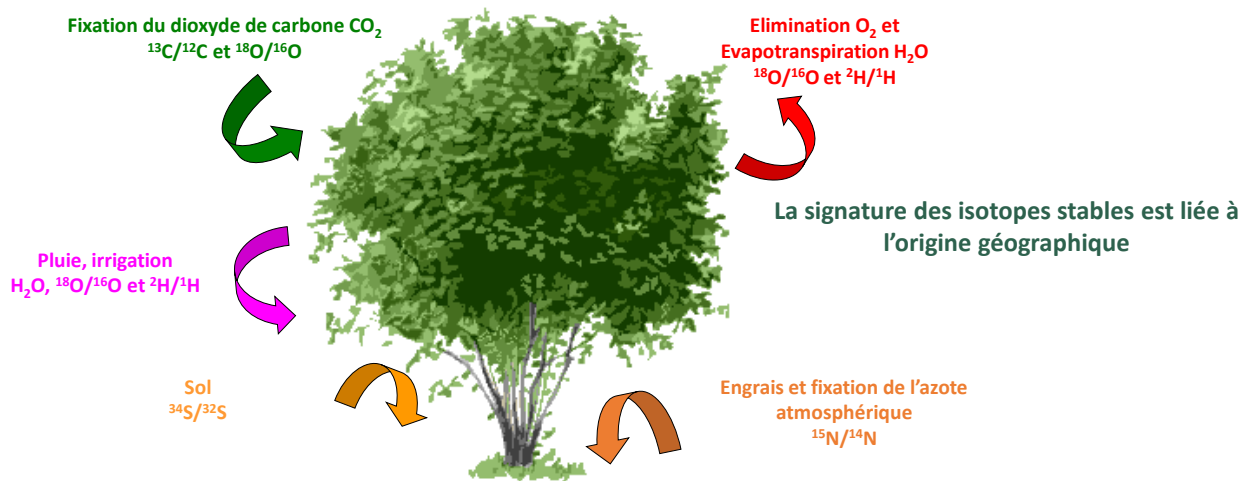
- ❑ **Analyse des isotopes radioactifs: détermination de l'activité du ^{14}C**
- ❑ **Analyse des ratios isotopiques stables ($^{13}\text{C}/^{12}\text{C}$; $^2\text{H}/^1\text{H}$; $^{18}\text{O}/^{16}\text{O}$, $^{34}\text{S}/^{32}\text{S}$)**
 - L'évaluation du rapport isotopique stable conduit à distinguer les molécules synthétiques et naturelles
 - Les isotopes stables sont utilisés comme marqueurs géographiques et variétaux



41

ANALYSE DES ISOTOPES STABLES

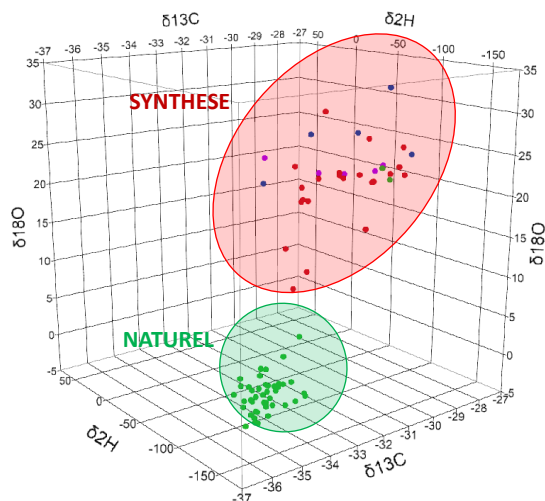
Harmonie entre le métabolisme de la plante et son environnement



42

EXEMPLE ETUDE DE L'HUILE ESSENTIELLE DE GAULTHÉRIE (*Gaultheria*)

Analyse isotopique globale par EA-C/P-IRMS



Identification d'ajout de salicylate de méthyle dans des huiles essentielles de wintergreen

- Différences isotopiques significatives entre les échantillons naturels et les échantillons de synthèse
- Discrimination possible avec les 3 éléments $\delta^{13}\text{C}$, $\delta^2\text{H}$ et $\delta^{18}\text{O}$

Légende

- HE Wintergreen sans impuretés (42 échantillons)
- HE Wintergreen avec impuretés (25 échantillons)
- Salicylate de méthyle commercial (5 échantillons)
- Salicylate de méthyle synthétisé au labo (5 échantillons)

Cuchet et al., 2019. Authentication of the naturalness of wintergreen (*Gaultheria* genus) essential oils by gas chromatography, isotope ratio mass spectrometry and radiocarbon assessment. Ind. Crops Prod. 142, 111873.

ETUDE DES HUILES ESSENTIELLES D'ALLIACÉES (*Allium*)

Problème de pureté



Huile essentielle
d'échalote



Huile essentielle
de poireau



Huile essentielle
de ciboulette



Huile essentielle
d'oignon

Composition de ces huiles essentielles similaires
Différences de prix importantes entre les origines botaniques

Comment différencier ces huiles essentielles ?
Discrimination botanique par analyse isotopique multi-éléments multi-composés

Cuchet et al., 2021. Multi-element (^{13}C , ^2H and ^{34}S) bulk and compound-specific stable isotope analysis for authentication of *Allium* species essential oils. Food Control. 126, 108086.

44

ETUDE DES HUILES ESSENTIELLES D'ALLIACÉES (*Allium*)

Problème de pureté



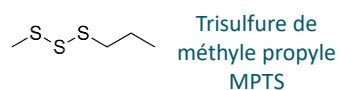
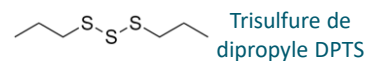
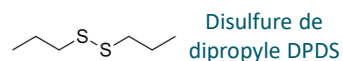
Famille des *Allium*

Oignons, poireaux, échalotes, ciboulettes

Extraction des huiles essentielles d'alliacées
Utilisation en agroalimentaire



Composition de l'huile essentielle
complexe avec de nombreux
organosulfures



Mesures isotopiques ciblées des principaux
métabolites organosulfurés
3 ratios isotopiques d'intérêt: $\delta^{13}\text{C}$, $\delta^2\text{H}$ et $\delta^{34}\text{S}$



Cuchet et al., 2020. Apport de l'analyse multi-isotopes pour l'authentification des huiles essentielles.. Annales des falsifications, de l'expertise chimique et toxicologique. 992, 58-70.

45

LDA Scores (axes F1 and F2: 94,49 %)

LDA Loadings (axes F1 and F2: 94,49 %)

Legend:

- Chive
- Shallot
- Onion
- Leek

Cuchet et al., **2021**. Multi-element (^{13}C , ^2H and ^{34}S) bulk and compound-specific stable isotope analysis for authentication of *Allium* species essential oils. Food Control. 126, 108086.

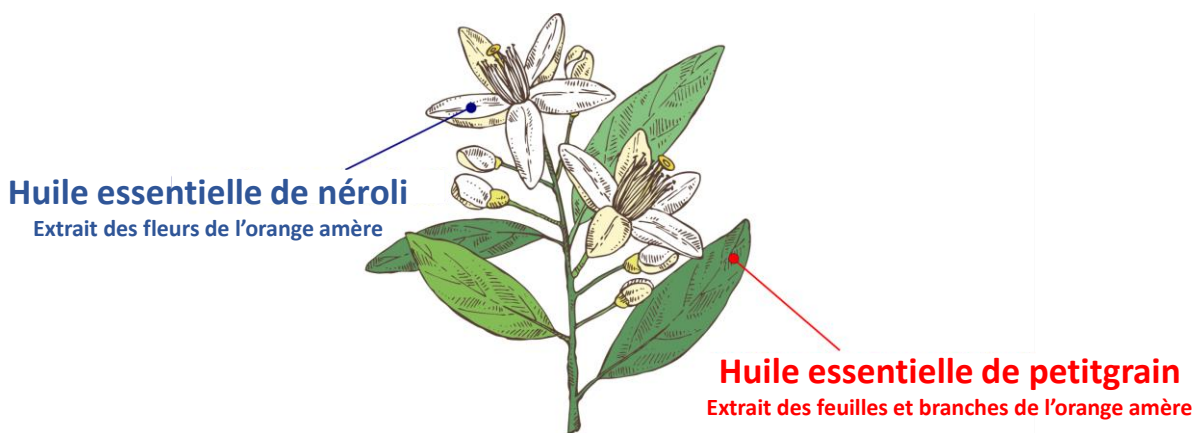
Analyse discriminante des Alliacés

Confusion matrix for the estimation sample:						
from \ to	Ciboulette	Echalote	Oignon	Poireau	Total	% correct
Ciboulette	7	0	0	0	7	100,00%
Echalote	0	4	0	0	4	100,00%
Oignon	0	0	5	0	5	100,00%
Poireau	0	0	0	6	6	100,00%
Total	7	4	5	6	22	100,00%
Confusion matrix for the validation sample:						
from \ to	Ciboulette	Echalote	Oignon	Poireau	Total	% correct
Ciboulette	5	0	0	0	5	100,00%
Echalote	1	4	0	0	5	80,00%
Oignon	0	0	8	0	8	100,00%
Poireau	0	0	0	2	2	100,00%
Total	6	4	8	2	20	95,00%

47

ETUDE DE L'HUILE ESSENTIELLE DE NÉROLI (*Citrus aurantium* L.)

Problème de pureté

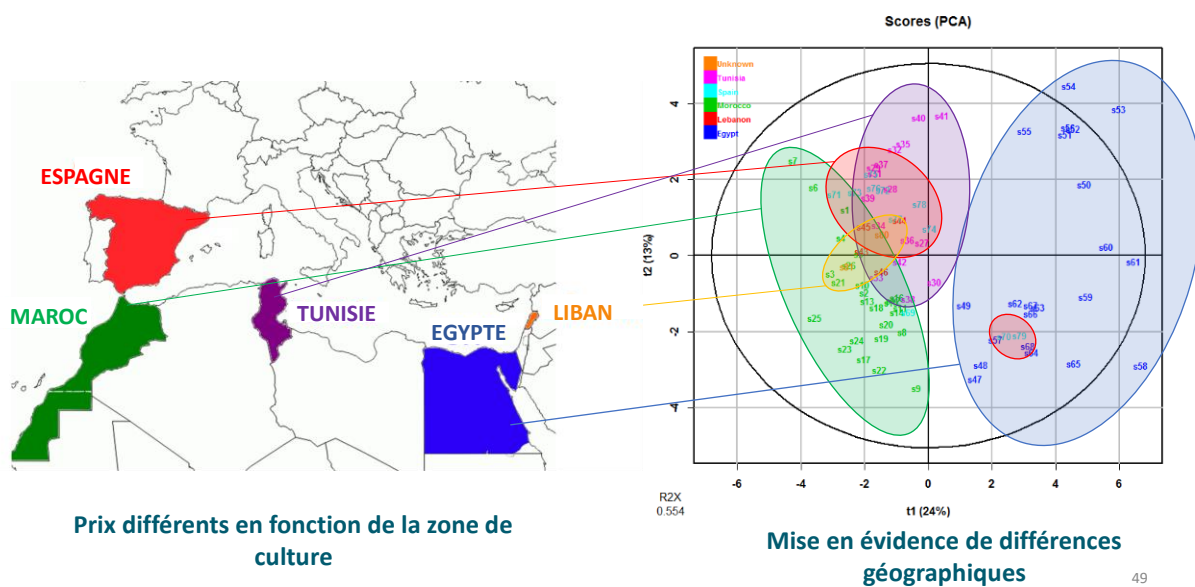


Compositions des deux huiles essentielles très proches
Aucun marqueur ne permet de distinguer un ajout de petitgrain dans le néroli

48

ETUDE DE L'HUILE ESSENTIELLE DE NÉROLI (*Citrus aurantium* L.)

Discrimination géographique



Cuchet et al., 2021. Determination of enantiomeric and stable isotope ratio fingerprints of active secondary metabolites in neroli (*Citrus aurantium* L.) essential oils for geographical and purity authentication. Soumis à publication

ETUDE DE L'HUILE ESSENTIELLE DE NÉROLI (*Citrus aurantium L.*)

Discrimination botanique

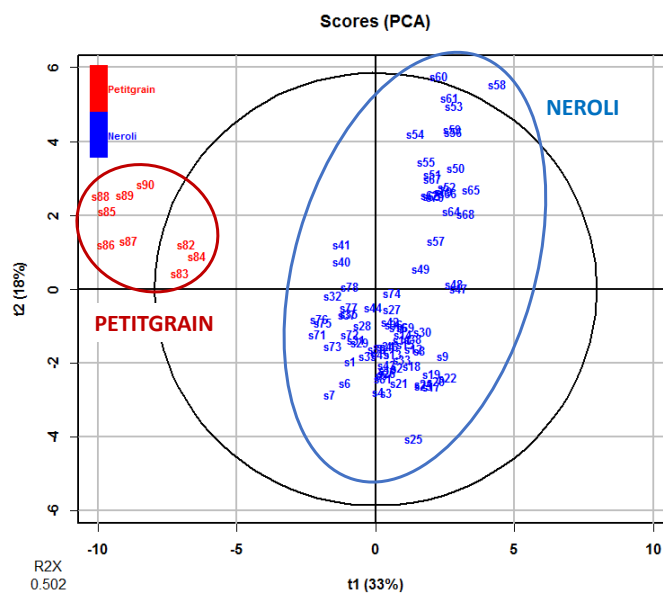
Etude de 90 échantillons de néroli et de petitgrain

Nombreux composés chiraux

Analyse énantiomérique de 14 molécules chirales

Nombreuses molécules en concentration significative

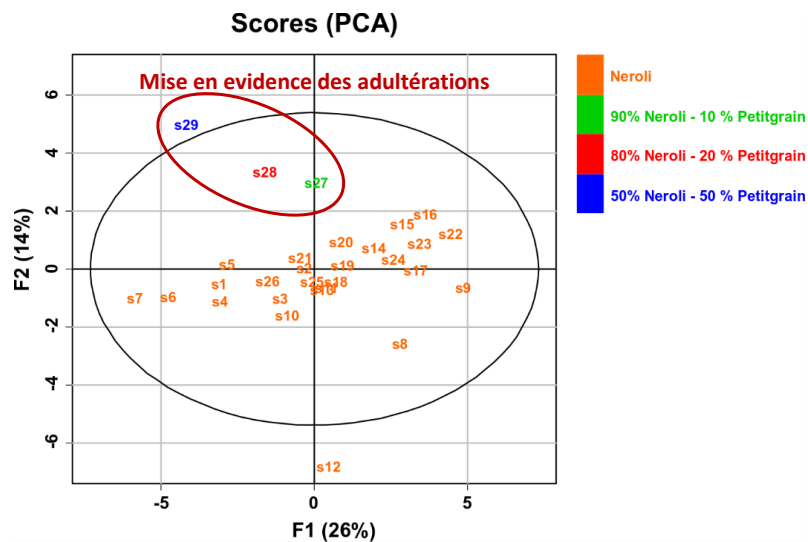
Analyse isotopique $\delta^{13}\text{C}$ et $\delta^2\text{H}$ de 10 molécules



Cuchet et al., 2021. Determination of enantiomeric and stable isotope ratio fingerprints of active secondary metabolites in neroli (*Citrus aurantium L.*) essential oils for geographical and purity authentication. Soumis à publication

ETUDE DE L'HUILE ESSENTIELLE DE NÉROLI (*Citrus aurantium L.*)

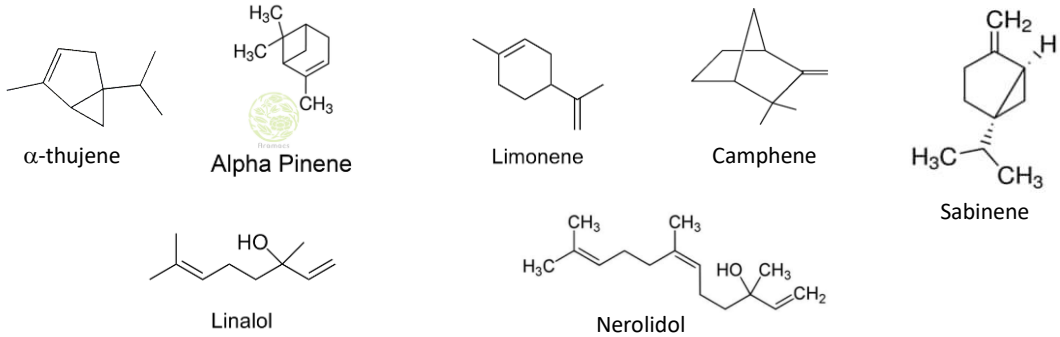
Adultération d'un échantillon de néroli, origine Maroc



Cuchet et al., 2021. Determination of enantiomeric and stable isotope ratio fingerprints of active secondary metabolites in neroli (*Citrus aurantium L.*) essential oils for geographical and purity authentication. Soumis à publication

ETUDE DE L'HUILE ESSENTIELLE DE NÉROLI

- **Analyses énantiosélectives**



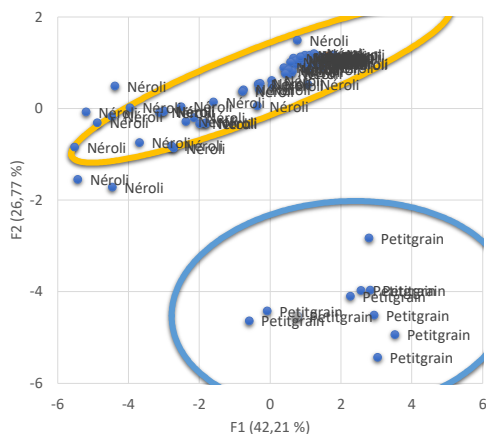
R-(-) (%)	S-(-) (%)	1S, 4R-(-) (%)	S-(-) (%)	S-(-) (%)	R-(+) (%)
α -Thujene	α -Pinene	Camphene	Sabinene	β -Pinene	Limonene
R-(-) (%)	R-(-) (%)	R-(+) (%)	R-(-) (%)	(S) (%)	
Linalol	Terpinene-4-ol	α -Terpineol	Acetate de linalyle	(E)-Nerolidol	

ETUDE DE L'HUILE ESSENTIELLE DE NÉROLI (*Citrus aurantium* L.)

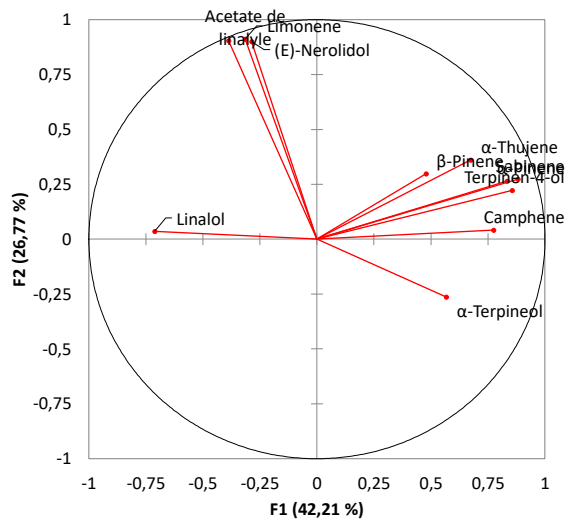
Discrimination géographique

Séparation
NEROLI/PETITGRAIN

Observations (axes F1 and F2:
68,98 %)



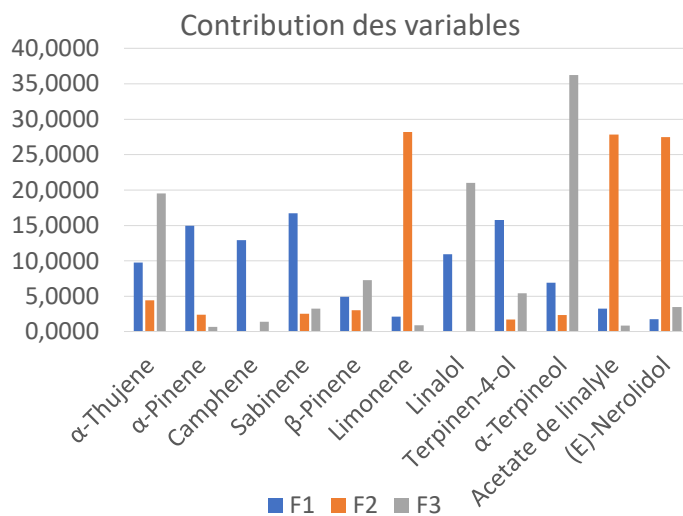
Variables (axes F1 and F2: 68,98 %)



ETUDE DE L'HUILE ESSENTIELLE DE NÉROLI (*Citrus aurantium L.*)

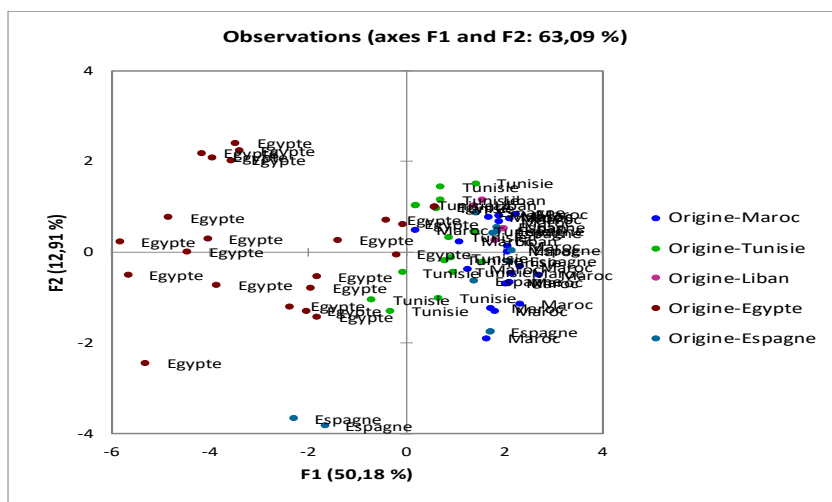
Discrimination géographique

Séparation
NEROLI/PETIGRAIN



ETUDE DE L'HUILE ESSENTIELLE DE NÉROLI (*Citrus aurantium* L.)

Discrimination géographique



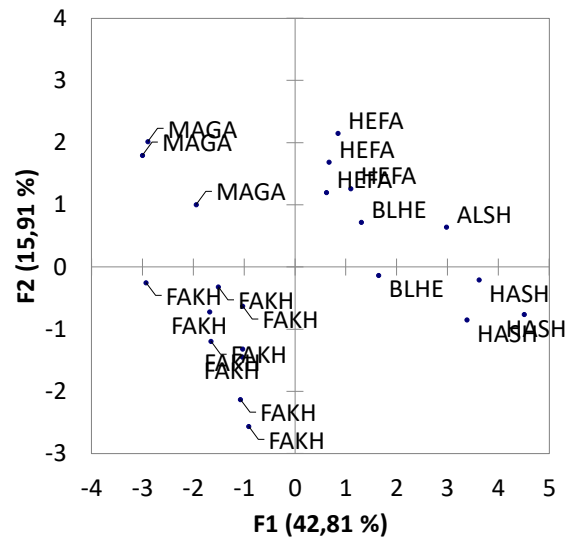
ETUDE DE L'HUILE ESSENTIELLE DE NÉROLI (*Citrus aurantium L.*)
Discrimination géographique

Confusion matrix for the estimation sample:							
from \ to	Egypte	Espagne	Liban	Maroc	Tunisie	Total	% correct
Egypte	14	0	0	0	0	14	100,00%
Espagne	0	5	0	0	0	5	100,00%
Liban	0	0	3	0	0	3	100,00%
Maroc	0	0	0	15	0	15	100,00%
Tunisie	0	0	0	0	9	9	100,00%
Total	14	5	3	15	9	46	100,00%

Confusion matrix for the validation sample:							
from \ to	Egypte	Espagne	Liban	Maroc	Tunisie	Total	% correct
Egypte	8	0	0	0	0	8	100,00%
Espagne	2	0	0	3	0	5	0,00%
Liban	1	0	0	0	0	1	0,00%
Maroc	0	0	0	5	0	5	100,00%
Tunisie	3	0	0	3	0	6	0,00%
Total	14	0	0	11	0	25	52,00%

ETUDE DE L'HUILE ESSENTIELLE DE NÉROLI (*Citrus aurantium* L.)

Répartition des Fournisseurs NEROLI Egyptien



CONCLUSIONS & PERSPECTIVES

59

CONCLUSIONS

Approche Multianalytique

- Efficacité de l'outil analytique pour distinguer les multiples adultérations possibles: Discrimination variétale et géographique, naturalité, ...
- Complémentarité des techniques pour mettre en place un arbre de décision et améliorer le diagnostic final
- D'autres techniques sont efficaces: IR, SNIF-NMR, GC*GC, HPTLC, ...

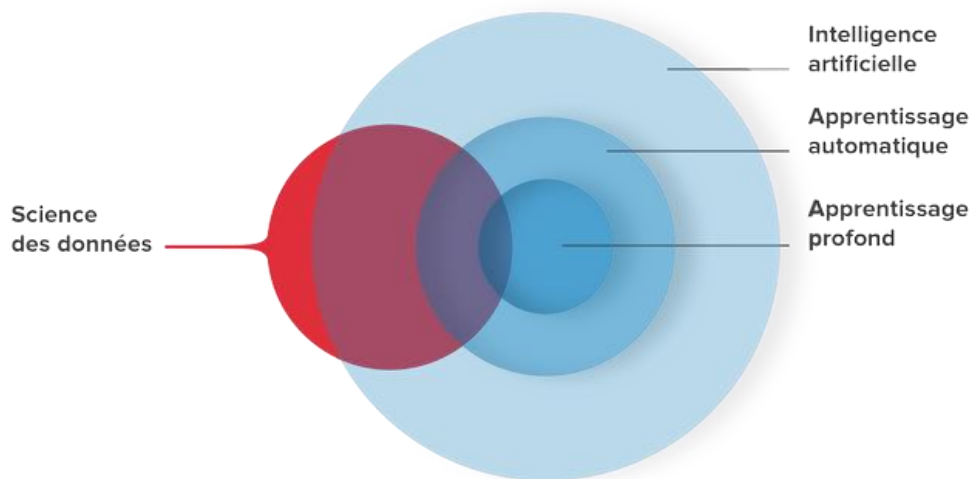
Etablissement de banques de données

- Construction de bases de données afin d'établir des critères de conformités
- Permet de rendre un verdict cohérent sur l'authenticité des matières premières.

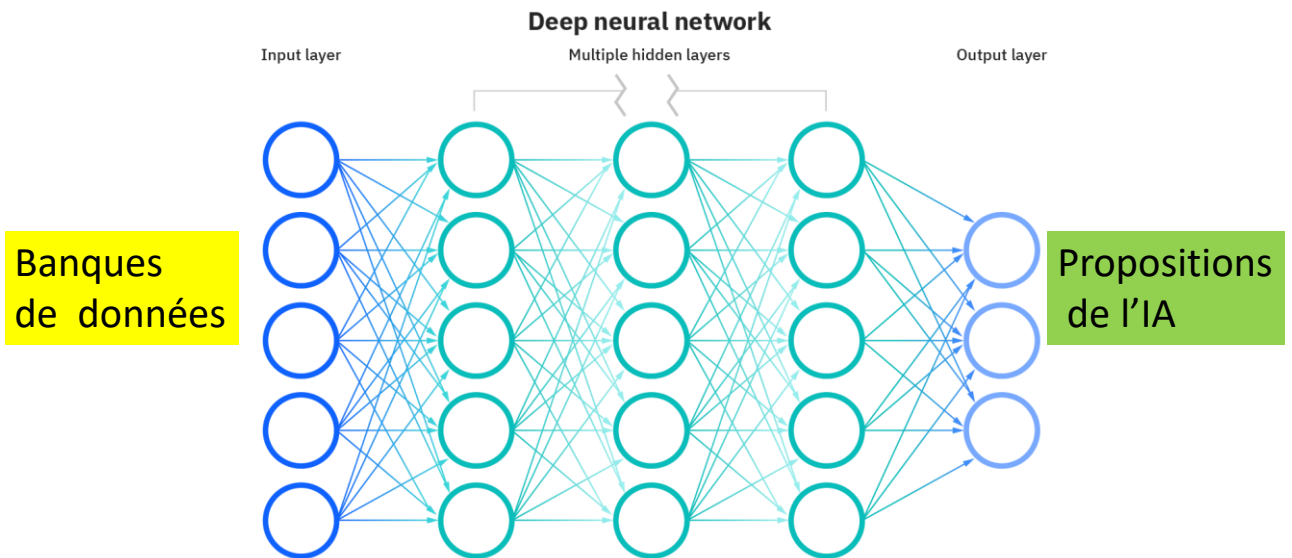
60

CONCLUSIONS

La construction de **bases de données fiables et cohérentes** est une des composantes essentielles pour la mise en œuvre **de procédés d'intelligence artificielle ...**



PERSPECTIVES



A suivre ...

63