



N° d'ordre : 57/ 2014

UNIVERSITE SULTAN MOULAY SLIMANE

Faculté des Sciences et Techniques

Béni-Mellal



UFR « Géosciences et Ressources Naturelles »

THÈSE

Présentée par

EL ABADI FATIHA

Pour obtenir le grade du

Doctorat National

Spécialité : Chimie Appliquée

**Application de la spectrophotométrie Infrarouge et de la Chimiométrie
pour le contrôle qualité des huiles alimentaires « huiles d'olive et
d'argan »**

Soutenue le 11 /10/2014 devant la commission d'examen :

| | | | |
|--------------------|---------------------|------------|------------------------|
| Président | A. JAOUAD | PES | FS-Marrakech |
| Rapporteurs | D. CHARQAoui | PES | FS-Marrakech |
| | A. ANOUAR | PES | FST-Settat |
| Examineurs | F. KZAIBER | PES | FST-Béni Mellal |
| | A. HAFID | PES | FST-Béni Mellal |
| | A. HASIB | PES | FST-Béni Mellal |
| | A. OUSSAMA | PES | FST-Béni Mellal |

Dédicaces

Je dédie ce travail

À celui qui m'a entouré de sa générosité divine et sa clémence éternelle.

Dieu tout puissant.

À mes chers parents, cœur de ma vie

Vous avez été avec moi chaque moment cœur et âme, vous avez veillé sur moi jour et nuit, de loin ou de près, vous avez toujours éclairé ma voie par vos conseils et votre vertu. Rien au monde ne pourrait compenser tous les sacrifices que vous avez consentis pour mon éducation et mon bien être. Que Dieu puisse vous accorder une longue vie pleine d'amour, de bonheur et de paix.

À mes frères, mes sœurs, et que ce travail soit le témoignage de mon amour fraternel. Je vous souhaite le bonheur et le succès.

À ma famille.

En témoignage de mon attachement

À tous mes professeurs

À tous mes amies (s).

À tous ceux qui me sont chers (es).

Remerciements

Ce travail de thèse a été réalisé au sein du Laboratoire de Spectro-Chimimétrie Appliquée et Environnement (LSCAE), Faculté des Sciences et Techniques, Université Sultan Moulay Slimane.

Même si parfois les mots semblent fades à côté de la profondeur des sentiments, il faut pourtant les concrétiser en remerciements, pour honorer tous ceux qui m'ont aidée à franchir ce pas vers l'avenir.

Je tiens à remercier mon encadrant le professeur **ABEDELKHALEK OUSSAMA** qui a bien voulu m'accueillir dans son laboratoire et qui m'a accompagné et guidé tout au long de ce travail. Je le remercie sincèrement pour son sens critique, sa présence au moment de doute, son professionnalisme, ses instructions claires et ses encouragements, pour la confiance qu'elle m'a accordée, pour son soutien, ses conseils qui m'ont permis d'évoluer dans ma vision de la recherche et dans la façon de la mener.

Je voudrais remercier le professeur **ELBAGHDADI MOHAMED** Directeur de L'UFR Géoscience et Ressources Naturelles.

Je remercie très sincèrement le professeur **RAUMA TAULER**, qui m'a accueilli dans son équipe de recherche, IDAEA-CSIC de Barcelone et m'a permis d'effectuer un stage dans son laboratoire et m'avoir consacré du temps afin de me faire profiter de leurs connaissances et de leurs expériences considérables en chimimétrie.

J'adresse un merci très particulier au professeur **ANNA DE JUAN CAPDUVILLA**, du département de chimie analytique Université de Barcelone pour m'avoir accueilli au sein de son laboratoire, et m'a co-encadré avec une grande attention pendant mon séjour. Je le remercie vivement pour ses conseils, pour son soutien amical.

J'exprime aussi ma profonde gratitude au professeur **OLIVIER DEVOS** maître de conférence au Laboratoire de Spectrochimie Infrarouge et Raman de Lille pour son aide, sa disponibilité, sa patience et ses conseils.

Je tiens aussi à remercier le professeur **ABDERRAFIA HAFID** qui m'a aidé à corriger ce manuscrit et pour ces encouragements.

Mes plus vifs remerciements vont au professeur **FOUZIA KZAIBER** pour ces encouragements et ses conseils.

Je remercie également Monsieur le professeur **JAOUAD ABDERRAHIM**, qui a très gentiment accepté de participer et pour l'honneur qu'il me fait en présidant le jury de cette thèse.

Je suis également très reconnaissant aux professeurs **DRISS CHERQAOUI** et **ABDELLAH ANOUAR** pour l'intérêt qu'elles ont bien voulu porter en acceptant de juger ce travail comme rapporteurs à la lumière de leur compétence notable.

Ma reconnaissance est adressée aux Professeurs **AZIZ HASSIB, ABDERRAFIA HAFID et FOUZIA KZAIBER** d'avoir accepté de juger ce travail en tant qu'examineurs.

Merci à l'ensemble de l'équipe du Laboratoire de Spectro-chimimétrie Appliquée et Environnement.

Merci aux thésards et tout particulièrement mes collègues, **WAFI TEROUZI** et **MAHFOUD ELBASSBASI**, ils m'ont fait vivre cette expérience entière dans une très bonne ambiance et certains n'ont pas hésité à placer leur confiance en moi.

Finalement, je remercie tous ceux ou celles qui ont agi dans l'ombre et participé discrètement à l'accomplissement de ce mémoire.

Résumé

L'huile d'olive représente l'un des produits stratégiques de la région euro-méditerranéenne. À l'instar des pays du bassin méditerranéen, le Maroc doit se plier aux exigences agronomiques et technologiques de son oléiculture.

L'arganier (*Argania spinosa*, Skeels L., Sapotacées) est un arbre endémique du sud-ouest marocain. En raison de la valeur élevée de ses huiles, la possibilité des pratiques illégales devrait être contrôlée.

La spectroscopie infrarouge à Transformée de Fourier (IRTF) est une technique puissante, et elle représente un outil d'analyse rapide et non destructive pour l'étude des huiles et graisses alimentaires, nécessitant peu ou pas de préparation des échantillons.

L'objectif de ce travail est d'explorer les potentialités de la spectroscopie IR et les méthodes chimiométriques pour prédire la qualité des huiles végétales

Dans ce travail, nous avons développé une nouvelle application basée sur la spectroscopie IRTF combinée avec des méthodes supervisées et des méthodes non-supervisées pour la classification des amandes d'argan par l'analyse hiérarchique ascendante. CHA a montré trois groupes distincts liés aux trois régions marocaines et a montré l'utilité de la spectroscopie infrarouge pour distinguer les origines des amandes d'argan. La PLS2 a été appliquée pour prédire l'origine géographique des amandes. L'autre partie de ce travail, elle a évalué l'efficacité de la spectroscopie infrarouge à transformée de Fourier (IRTF) à détecter l'adultération. Le premier axe porte sur la détection de l'adultération d'huile d'olive par l'huile de soja et l'huile de tournesol. La régression PLS et l'analyse en composante principale ACP ont été utilisées comme un outil analytique. Le deuxième axe porte sur la détection et la classification de l'adultération de l'huile d'argan par l'huile de soja et l'huile de tournesol.

Mots-clés : Spectroscopie Infrarouge, Chimiométrie, PLS, PCA, analyse discriminante, CPG, ATR, huile, FTIR, cluster analyses, Origine géographique, Amande d'argan

Abstract

Olive oil is one of the strategic products of the Euro-Mediterranean region. Like the Mediterranean countries, Morocco must comply with agronomic and technological requirements of the olive.

The argan tree (*Argania spinosa* L. Skeels, Sapotaceae) is an endemic tree of southwestern Morocco. Due to the high value of the product, the possibility of illegal practices should be controlled.

Fourier transform mid-infrared (FTIR) is a powerful technique, a tool for rapid and non-destructive analysis for the study of edible oils and fats, requiring little or no sample preparation.

The objective of this work is to explore the potential of IR spectroscopy and chemometrics methods to predict the quality of vegetable oils.

In the first part of this work, we have developed a novel technique based on FTIR spectroscopy combined with supervised methods and unsupervised methods applied for the classification of argan kernels by Hierarchical Cluster Analysis. HCA showed three distinct clusters related to the Moroccan regions and suggested the usefulness of IR spectroscopy to distinguish among kernel origins. The PLS2 was applied to predict the geographical origin of almonds.

The second part was directed towards the determination of fatty acids in oils by gas chromatography combined of classification tool to classify mixtures compared to pure oils.

The third part of this work is divided into two parts, has evaluated the effectiveness of Fourier transform mid-infrared (FTIR) to detect adulteration. The first part focused on the detection of adulteration of olive oil with soybean oil and sunflower oil; PLS regression and principal component analysis ACP were used as an analytical tool. The second axis focused on the detection and classification of the adulteration of argan oil by soybean oil and sunflower oil.

Keywords: infrared spectroscopy, chemometric, PLS, PCA, Discriminant analysis, CPG, ATR, oil, FTIR, cluster analysis , geographical origin, argan kernels

Sommaire

| | |
|---|----|
| INTRODUCTION GENERALE | 1 |
| ETUDE BIBLIOGRAPHIQUE | |
| Chapitre 1: ETUDES ET CARACTERISATIONS DES HUILES VEGETALES (HUILE D'ARGANE ET L'HUILE D'OLIVE) | 5 |
| A. L'ARGANIER..... | 5 |
| 1. Identification de l'huile d'Argan et du contexte général de sa production | 5 |
| 2. Caractéristiques physico-chimiques et usage de l'huile d'argan et l'arganier | 8 |
| Références | 12 |
| B.L'HUILE D'OLIVE | 15 |
| 1 .Généralités..... | 15 |
| 2 .Composition chimique des huiles végétales utilisées et facteurs influant la qualité de l'huile d'olive | 17 |
| Références | 22 |
| Chapitre 2: DETERMINATION DES ACIDES GRAS DES HUILES PAR CHROMATOGRAPHIE EN PHASE GASEUSE | 26 |
| 1. Réaction de transestérification des huiles..... | 26 |
| 1.1. Introduction..... | 26 |
| 1.2. Les aspects thermodynamiques | 27 |
| 2. Les différentes techniques d'analyse d'un mélange complexe | 28 |
| 2.1. La Chromatographie en Phase Gazeuse (CPG) | 28 |
| 2.2. Application de la chromatographie en phase gazeuse..... | 29 |
| Références | 32 |
| Chapitre 3 :L'INFRAROUGE ET SES APPLICATIONS DANS LES HUILES | 34 |
| 1. Généralité..... | 34 |
| 2. La spectroscopie infrarouge..... | 34 |
| 3. La Spectroscopie Moyen infrarouge (MIR) et ses applications | 38 |
| 3.1 Application de la spectroscopie Moyen Infrarouge | 39 |

| | |
|---|----|
| 3.2 Application de la spectroscopie moyenne infrarouge à l'étude des huiles..... | 40 |
| 3.3. Attribution des bandes d'absorption des lipides dans le moyen infrarouge..... | 40 |
| 4. La spectroscopie proche infrarouge et ses applications | 41 |
| Références | 43 |
| Chapitre 4 : LA CHIMIOMETRIE ET SES APPLICATIONS | 49 |
| 1. Définition de la Chimiométrie..... | 49 |
| 2. Pré traitements des données spectrales | 49 |
| 3. Méthodes descriptives ou exploratoires..... | 50 |
| 3.1. Analyse en composantes principales (ACP) | 50 |
| 3.2. Exemples d'applications de l'Analyse en Composantes Principales | 51 |
| 4. Les méthodes de classement..... | 51 |
| 4.1 Classification ascendante hiérarchique. | 52 |
| 4.2.CLV (Clustering of variables around Latent Variables)..... | 52 |
| 5. Méthodes prédictives | 53 |
| 5.1. Régression linéaire et Méthode des moindres carré PLS..... | 53 |
| 5. 2. ipls..... | 54 |
| 5.3.Exemples d'applications de la méthode PLS en analyse quantitative..... | 55 |
| 5.4. Analyse discriminante par les moindres carrés partiels (Partial Least Square Discriminant Analysis - PLSDA) | 56 |
| 6. Validation croisée ou cross validation..... | 58 |
| Références | 59 |
| PARTIE EXPERIMENTALE | |
| Chapitre 5 : CLASSIFICATION DES AMANDES D'ARGAN SELON L'ORIGINE GEOGRAPHIQUE PAR LA MOYEN INFRAROUGE COUPLEE A LA CHIMIOMETRIE | |
| 1. Introduction..... | 66 |
| 2. Matériel et méthodes..... | 67 |
| 3. Résultats et discussions..... | 71 |
| 3.1. Interprétation des prétraitements des données spectrales..... | 71 |
| 3.2. La classification hiérarchique ascendante (HCA) : Dendrogramme..... | 73 |

| | |
|---|-----------|
| 3.3. Classification PLS-DA et prédiction | 74 |
| Conclusion | 82 |
| Références | 83 |
| | |
| Chapitre 6 : CARACTERISATION DES HUILES PAR CHROMATOGRAPHIE EN PHASE GAZEUSE | 86 |
| 1. Introduction..... | 86 |
| 2. Matériels et méthodes | 86 |
| 2.1. Acides gras de références..... | 86 |
| 2.2. Principe | 87 |
| 2.3. Protocole expérimental utilisé pour la préparation et analyse des esters méthyliques d'acides gras..... | 88 |
| 2.4 Analyse par chromatographie en phase gazeuse des esters méthyliques d'acides gras..... | 88 |
| 3. Résultats et discussions..... | 89 |
| 3.1. Utilisation de l'analyse des acides gras..... | 89 |
| 3.2. Composition en acides gras des huiles..... | 90 |
| 3.3. Analyse en composantes principales..... | 93 |
| | |
| Conclusion | 95 |
| Références | 97 |
| | |
| Chapitre 7 : APPLICATION DE L'INFRAROUGE ET LES METHODES CHIMIOMETRIQUES POUR LA DETECTION DE L'ADULTERATION DES HUILES..... | 99 |
| | |
| Partie A :Détection de l'adultération de l'huile d'olive par l'huile de soja et l'huile de tournesol | 99 |
| 1. Introduction..... | 99 |
| 2. Matériel et méthode..... | 100 |
| 3. Résultats et discussions..... | 102 |
| 3.1. Interprétation des données spectrales..... | 102 |
| 3.2. Analyse en composante Principale | 105 |
| 3.3. Partial Least Square Régression..... | 106 |
| 3.4. Validation du modèle..... | 109 |
| 3.5. L'importance variable en projection ou VIP score..... | 110 |

| | |
|--|------------|
| Conclusion | 111 |
| Références | 113 |
| Partie B : Détection de l'adultération de l'huile d'argan par l'huile de soja | 115 |
| 1. Introduction..... | 115 |
| 2. Matériel et méthodes..... | 116 |
| 3. Résultats et discussions..... | 117 |
| Conclusion | 124 |
| Références | 126 |
| CONCLUSION GÉNÉRALE | 127 |

ANNEXE 1: Liste des publications

Liste des abréviations

AG : Acides gras

HCA : Analyse en Classification Hiérarchique

ACP: Analyse en Composantes Principales

PLS-DA: Analyse discriminante par les moindres carrés partiels (Partial Least Square Discriminant Analysis)

ANOVA : Analyse de variance «Analysis of variance»

CHA: Classification Hiérarchique Ascendante

HPLC : Chromatographie liquide à haute performance

CPG : Chromatographie en phase gazeuse

GC/MS : Chromatographie en phase gazeuse couplée à la spectrométrie de masse

CNUCED :(Conférence des Nations Unies sur le Commerce et le Développement

COI : Conseil oléicole international

RMSEC : Erreur quadratique moyenne d'étalonnage ("Root Mean Square Error of Calibration").

RMSECV : Erreur quadratique moyenne de validation croisée ("Root Mean Square Error of Cross-Validation").

H.A : Huile artisanale

EVOO : Huile d'olive extra vierge

H .P : Huile de presse

HPT: Huile de presse torréfiée

HPNT : Huile de presse non torréfiée

HLT : Huile de laboratoire torréfiée

HLNT : Huile de laboratoire non torréfiée

IR : Infrarouge

FTIR : Infrarouge à transformée de Fourier «Fourier transform infrared spectroscopy»

MIR : Moyen Infrarouge « Mid-infrared»

MSC : «Multiplicative scatter correction»

NIST: National Institute of Standards and Technology

NIPALS : «Nonlinear iterative partial least squares»

NIR : Proche Infrarouge «Near-infrared»

ATR : Réflexion totale atténuée «Attenuated total reflection»

PLS: Régression par les moindres carrés partiels (Partial Least Square)

RMN : Résonance magnétique nucléaire

SNV : Standard Normal Variate

TG : Triglycéride

UV : Ultraviolet

CV : Validation croisée «Cross-validation»

LV : Variable latente « Latent variable »

Liste des figures

| | |
|--|----|
| Figure 1: Carte de l'aire de répartition de l'arganier au Sud-Ouest marocain | 7 |
| Figure 2 : Evolution de la production mondiale de l'huile d'olive en 2011 | 16 |
| Figure 3 : Principaux pays consommateurs en 2011/2012 en tonnes - source COI - graph. A.Paris/Afidol) | 17 |
| Figure 4. Composition de l'olive (Nefzaoui, 1984) | 20 |
| Figure 5 : Réaction générale de transestérification de triglycérides avec un alcool (R-OH) avec R ₁ , R ₂ et R ₃ des groupements alkyles présents dans les acides gras | 26 |
| Figure 6 : Les trois étapes successives et réversibles de la réaction de transestérification | 27 |
| Figure 7 : Le spectre électromagnétique, d'après Dalibart et Sevant [5] | 35 |
| Figure 8: Schéma de principe d'un spectromètre infrarouge à transformée de Fourier [8] | 36 |
| Figure 9 : Principe de la réflexion totale atténuée (ATR) sur diamant | 38 |
| Figure 10 : Spectre moyen infrarouge pour les amandes d'argan pour les régions Agadir (a), Ait baha(b) et Guelmum(c) respectivement | 72 |
| Figure 11 : Dendrogramme avec la distance euclidienne | 74 |
| Figure 12: Score plot pour le modèle PLS-DA des amandes argans pour les trois régions Ait baha (AITB), Agadir (agt) and Guelmim(glm) avec le prétraitement premier dérivée | 77 |
| Figure 13: les valeurs prédites en fonction des valeurs mesurer pour Ait baha, Agadir et Golmume avec le modèle PLS -DA | 79 |
| Figure 14 : La Prédiction de validation externe des amandes d'argan : (a) classe Ait baha; (b) classe Agadir et (c) la classe Guelmim. | 81 |
| Figure 15 : Chromatogramme pour les différents mélanges de l'huile d'olive adultérée par l'huile de soja et les autres huiles végétales : 1acide palmitique, 2 acide linoléique 3, acide oléique, 4 acide stéarique | 90 |
| Figure 16 : Chromatogramme des acides gras pour les huiles végétales : C huile de soja , F huile d'olive : 1acide palmitique, 2 acide stéarique 3, acide oléique, 4 acide linoléique, 5 acide linoléique | 91 |
| Figure 17: Plot d'Analyse en composante principale avec le prétraitement centré | 93 |
| Figure 18: La variance expliquer en fonction des nombre de composantes principales | 94 |

| | |
|---|-----|
| Figure 19: Résultats de la 1 ^{ère} dérivée avec un polynôme égale à 2 et un smothing 3 | 94 |
| Figure 20 : Résultats de la variance résiduelle | 95 |
| Figure 21 : Spectre moyen infrarouge de l'huile d'olive | 102 |
| Figure 22 : Spectres MIR de l'huile d'olive (bleu), huile de soja (rose) et l'huile de tournesol (rouge) | 104 |
| Figure 23 : les trois spectres moyens infrarouges de l'huile d'olive de l'huile de soja et l'huile de tournesol purs. | 104 |
| Figure 24 a) score plot du 1 ^{ère} principal composant pour la classification des mélanges binaires huile de soja –huile d'olive et de huile de tournesol-huile d'olive ; b): 3D scores plot des trois premiers PC de la classification des mélanges binaires huile de soja –huile d'olive et de huile de tournesol-huile d'olive à l'étape de l'étalonnage; c) les résultats de la validation externe (Cross validation: leave one out). | 107 |
| Figure 25: les valeurs prédites par le model PLS en fonction des valeurs mesurer a) pour le modèle de calibration b) pour le model de validation externe | 109 |
| Figure 26 : Importance variable de la projection, les scores de VIP | 110 |
| Figure 27 : spectres infrarouges FTIR pour l'huile d'argan (bleu), huile de soja (rouge) et l'huile de tournesol (rose) | 117 |
| Figure 28 : Les spectres moyens infrarouges des échantillons de l'huile d'argan –huile de tournesol dans la région 3200-600 cm ⁻¹ | 117 |
| Figure 29 : les spectres de niveau différent d'argan avec de l'huile de soja dans la zone 700-1900 cm ⁻¹ . | 118 |
| Figure 30a: – Les concentrations prédites PLS en fonction des concentrations réelles de l'huile de tournesol dans l'huile d'argan dans l'étape de calibration (bleu) et dans l'étape de validation croisée (rouge)avec la première dérivée de type savitzky- golay | 121 |
| Figure 31b: Les concentrations prédites PLS en fonction des concentrations réelles de l'huile de soja dans l'huile d'argan dans l'étape de calibration (bleu) et dans l'étape de validation croisée (rouge)avec le prétraitement SNV +SCM | 121 |

Liste des tableaux

| | |
|---|-----|
| Tableau 1: Composition chimique de la pulpe (% m/m) d'après SANDRET (1956) | 9 |
| Tableau 2 : Composition des tourteaux à l'état frais (% m/m) d'après CHARROUF (1991) | 10 |
| Tableau 3 : Principales constantes physico-chimiques de l'huile d'argan vierge | 11 |
| Tableau 4 : Pourcentage des acides gras dans quelques huiles végétales (Harwood et Aparicio, 2000) [32]. | 21 |
| Tableau 5 : longueurs d'onde (μm) et nombres d'onde (cm^{-1}) pour les régions du proche, moyen et lointain infrarouge. | 35 |
| Tableau 6 : Les attributions chimiques des principales bandes d'absorption des lipides dans le MIR [49]. | 41 |
| Tableau 7 : Principaux travaux analytiques sur divers produits alimentaires, faisant intervenir la PCA ou la LDA ou les deux comme outils de traitement des données (de 1990 à 2002) | 51 |
| Tableau 8 : Les données climatiques des trois régions étudiées. | 68 |
| Tableau 9 : Les paramètres statistiques obtenus avec la validation croisée pour le modèle PLS-DA | 76 |
| Tableau 10: Les résultats de classification des échantillons de validation externe des amandes d'argan. | 80 |
| Tableau 11 : Acides gras estérifiés, saturés et insaturés (points de fusion tiré de : Handbook of Chem. & Phys.1968/69) | 87 |
| Tableau 12 : Conditions d'injections | 89 |
| Tableau 13: Résultat de la composition en acides gras des mélanges d'huile d'olive avec des huiles végétales | 92 |
| Tableau 14: le % de la variance pour le modèle ACP | 105 |
| Tableau 15 : les régions spectrales où les VIP ont été jugées sur le seuil de un. | 111 |
| Tableau 16 : les résultats de différents prétraitements ou combinaisons de traitements faites | 119 |
| Tableau 17 : Paramètres de calibration (cross validation) pour le modèle régression PLS de l'huile de soja | 120 |

| | |
|---|-----|
| Tableau 18 : Résultats de la validation externe pour le tournesol avec 5 CP | 122 |
| Tableau 19 : Les paramètres de prédiction avec le modèle PLS pour le tournesol | 123 |
| Tableau 20 : résultat de la validation externe pour le soja | 123 |
| Tableau 21 : le % de la variance pour le modèle ACP | 124 |

INTRODUCTION GÉNÉRALE

La qualité nutritionnelle et organoleptique des aliments est un important sujet puisqu'il commence sur le terrain par le suivi et le soutien scientifique et technique des filières de production (élevages, agriculture...) et se poursuit jusqu'au produit fini que nous retrouvons dans notre assiette.

L'huile d'olive est un élément clé du régime méditerranéen. Très présente dans l'alimentation des pays méditerranéens et préconisée par de nombreux diététiciens, elle a acquis une place essentielle dans la recherche capitalisant sur ses vertus aussi bien médicinales qu'alimentaires [1]. Elle est l'une des huiles végétales les plus anciennes et la seule qui peut être consommée sous sa forme brute sans traitement préalable [2].

L'arganier (*Argania spinosa*) est une espèce végétale endémique du sud-ouest Marocain, il est d'ailleurs le seul arbre à croître au Nord du Sahara où il couvre une surface de 820 000 ha selon les chiffres de la direction de développement forestier.

Cette huile est particulièrement riche en acide oléique (45% à 48%) et en acide linoléique (32% à 35%) et tout à fait dépourvu d'acides gras polyinsaturés de la famille w-3 [3]. Elle est actuellement utilisée dans la médecine traditionnelle pour ses propriétés cosmétique, bactéricide et fongicide. Cette huile reçoit une attention croissante en raison de son potentiel pharmacologique comme la réduction du risque cardio-vasculaire [4] et la présence de substances anti-cancéreuse [5].

Face à la concurrence, d'autres huiles alimentaires végétales (huile de palme, huile de soja, huile de colza, etc.), il est important de redonner à l'huile d'olive et l'huile d'argan les places qu'elles méritent sur les marchés nationaux et internationaux. Pour atteindre cet objectif, il est nécessaire de faire appel à de nouvelles méthodes d'analyse développées pour quantifier, détecter l'adultération de l'huile d'olive et l'huile d'argan par des huiles végétales moins chères.

Ces méthodes devront être rapides, précises et leurs limites de quantification devront être suffisamment faibles pour répondre aux différentes normes. Cependant les analyses chimiques classiques sont en général longues, nécessitent parfois l'utilisation de réactifs onéreux et polluants, et ne sont utilisables que par des opérateurs qualifiés. Une méthode spectrale appropriée, telle que la spectroscopie infrarouge, rend possible les analyses en quelques minutes.

La spectroscopie IR trouve aujourd'hui des applications dans de nombreux domaines : l'agro-alimentaire, la médecine, l'environnement, la biologie, la microbiologie, la chimie. Elle

peut donner des informations sur la composition des denrées agroalimentaires et sur la conformation des molécules biologiques, même très complexes [6].

Cette technique nécessite, le plus souvent, pour être exploitée de manière pertinente, le recours à des méthodes de traitement mathématiques et statistiques de données. La falsification de l'huile avec des huiles comestibles moins coûteuses est un enjeu majeur pour l'industrie de l'huile d'olive.

Avec l'essor de ces nouvelles techniques, de nouvelles approches mathématiques ou statistiques sont utilisées pour chercher à exploiter au mieux les informations contenues dans les spectres et à établir des protocoles facilement utilisables grâce à la chimiométrie. En effet, les spectres ne sont plus interprétés sous leur forme brute généralement entachés de défauts, mais subissent des traitements comme des dérivées premières et secondes [7], des décompositions spectrales [8].

Le premier chapitre de ce mémoire présente une synthèse bibliographique concernant l'huile d'argan et l'huile d'olive: leurs caractéristiques chimiques et physico-chimiques et les facteurs influençant leur qualité.

Le deuxième chapitre présente l'application des méthodes chromatographiques pour l'analyse et le contrôle qualité des huiles.

Le troisième et quatrième chapitre détaillent la spectroscopie infrarouge et la chimiométrie, le principe de la spectroscopie IR. Nous allons aborder quelques applications de la technique pour voir l'avantage de l'utilisation combinée des techniques spectroscopiques et les méthodes chimiométriques capables d'extraire et modéliser l'information présente dans les spectres.

Dans la partie expérimentale, le chapitre cinq est consacré à l'identification de l'origine géographique des amandes d'argan provenant de trois zones géographiques différentes. Cette étude consiste à développer une nouvelle application basée sur la spectroscopie FTIR combinée avec des méthodes supervisées et des méthodes non-supervisées comme outil rapide, peu coûteux et non destructives, utile pour déterminer l'origine géographique des amandes d'argan.

Le chapitre six vise la caractérisation des acides gras par la chromatographie en phase gazeuse des mélanges des huiles et leur classification au moyen de l'analyse en composantes principales (ACP).

Le dernier chapitre, divisé en deux parties, traite la capacité de l'infrarouge couplé à des méthodes chimiométriques à la détection de l'adultération des huiles. Nous avons utilisé

la spectroscopie infrarouge à transformée de Fourier (FTIR) pour classer les différentes huiles végétales adultérant l'huile d'olive vierge et l'huile d'argan. La régression PLS est utilisée pour construire un modèle d'étalonnage entre les concentrations réelles des huiles étudiées et les concentrations prédites. Nous avons également quantifié le taux d'adultération de l'huile d'olive vierge et huile d'argan avec les huiles végétales à des pourcentages massiques allant de 0 à 30 % pour les huiles de soja et de tournesol. Les données spectrales sont enregistrées dans la région moyen infrarouge de $4000 - 400 \text{ cm}^{-1}$ pour tous les échantillons puis analysées à l'aide de différents outils chimiométriques telles que l'analyse en composantes principales (ACP) et la Régression aux sens des moindres carrés partiels (PLS).

Références bibliographiques

- [1] Luaces, P. Pérez, A. G., et Sanz, C. 2003 .Role of olive seed in the biogenesis of virgin olive oil aroma. *Journal of Agricultural and Food Chemistry*, 51, 4741-4745.
- [2] Boscou, D. 1996 Olive Oil Composition. In *Olive Oil: Chemistry and Technology*. AOACS Press, USA, 52-83, 85-127.
- [3] Berrougui H, Ettaib A, Herrera Gonzalez MD, Alvarez de sotomayer 2004. Hypolipidemic and hypocholesterolemic effect of argan oil (*Argania spinosa*) lowers blood pressure and improves endothelial dysfunction in spontaneously hypertensive rats. *Br J Nutr*, 92, 921-9,
- [4] Keys A. 1995: Mediterranean diet and public health: personal reflections. *Am. J Clin Nutr*, 61 (suppl.), 1321-1323,
- [5] Khalloukhi F., Younos C., Soulimani R. Oster T., Charrouf Z., Spiegelhalder B., Bartsch H., Owen R.W 2003, Consumption of argan oil with its unique profile of fatty acids, tocopherols, squalene, sterols and phenolic compounds should confer valuable cancer chemopreventive effects. *Eur J Cancer Prev*, 12, 67-75,
- [6] Bertrand, D. Courcoux, P. et Qannari El, M. 2000 « Méthodes exploratoires » dans « la spectroscopie infrarouge et ses applications analytiques » Tec & Doc,Lavoisier. Paris, 261-266.
- [7] Navas N., Romero-Pastor J., Manzano E. et Cardell C. 2010. Raman spectroscopic discrimination of pigments and tempera paint model samples by principal component analysis on first-derivative spectra, *Journal of Raman Spectroscopy*, 41, 1486-1493
- [8] Shashoua Y., Berthelsen M.B.L.D. et Nielsen O.F. 2006. Raman and ATR-FTIR spectroscopies applied to the conservation of archaeological Baltic amber, *Journal of Raman Spectroscopy*, 37, 1221-1227

Chapitre 1 : ETUDE DES HUILES VEGETALES (HUILE D'ARGANE ET L'HUILE D'OLIVE)

A. L'ARGANIER

I. Identification de l'huile de l'Argan et du contexte général de sa production

I.1 L'arganier et sa répartition

L'arganier (*Argania spinosa* (L.) Skeels), est une espèce endémique du sud-ouest marocain. C'est le seul représentant de la famille des sapotacées (famille tropicale et subtropicale) au nord du Sahara en Afrique. De l'avis de plusieurs auteurs, il est l'arbre le plus original et le plus remarquable de l'Afrique du Nord. C'est dû essentiellement à son intérêt botanique et taxonomique et l'usage multiple de ses produits (feuillage, bois, fruits, écorce,...).

L'arganier est un arbre épineux à port d'olive. Il forme des arbres à troncs tortueux pouvant atteindre 8 à 10 mètres en fonction des situations écologiques. La cime grande et étalée, le tronc vigoureux, l'écorce craquelée et les ramifications sont très denses avec de jeunes rameaux épineux avec des fruits en forme de baies.

Les phéniciens ont connu l'arbre et ont utilisé son huile [1].

Les études édaphiques ont montré que l'arganier est un arbre très résistant qui peut vivre de 150 à 250 ans. Il est aussi indifférent à la nature physico-chimique du substrat.



Photo 1 : Arbre d'argan.

L'arganier caractéristique du Sud-Ouest Marocain, couvre actuellement une surface de 820 000 ha selon les chiffres de la direction de développement forestier, soit 14,3 % de la

superficie totale de la forêt naturelle au Maroc. Sa superficie actuelle le place parmi les toutes premières essences forestières du Maroc où il n'est devancé que par le chêne vert (*Quercus rotundifolia L.*) avec 1.364.1000 ha [2]. Les principales Arganraies sont celles de Tamanar, Ida ou Trhouma, Admine, Mesguine..., on entend par l'appellation arganraies, les écosystèmes formés par l'arganier en tant qu'espèce dominante. En considérant une densité de 20 arbres/ha, El Aich différencie entre l'aire de l'arganier plus vaste que la forêt d'arganiers [3]. Cette aire couvre les provinces de Safi, Essaouira, Agadir, Tiznit et Guelmim. D'autres essences se trouvent isolées dans le nord-Est du pays, dans les massifs montagneux de Béni-Snassen et dans la vallée de l'oued Grou au sud-Est de Rabat.

L'arganier peuple des bioclimats sahariens, arides et semi arides du Sud-Ouest marocain. Il occupe la tranche altitudinale allant du bord de la mer à 1600-1700 m d'altitude sur le versant sud du haut Atlas occidental et l'anti-atlas.

L'arganier supporte une pluviométrie très variable allant de 120 mm (Guelmim) à 500 mm (les reliefs des Ida-Outanane), [4]. Cet écart des précipitations dont l'arganier s'accommode est suffisamment large pour qu'il soit considéré comme résistant à la sécheresse et extrêmement peu exigeant en eau.

Température : l'arganier supporte un intervalle de température très large allant de 3°C (l'isotherme 3°C, freine son développement dans la partie Nord en altitude) à 50°C (enregistrée à Taroudant) [5].

Le facteur édaphique ne joue aucun rôle dans le développement de l'arganier puisqu'il croit sur tous les sols (siliceux, calcaires, schisteux ou alluvions), mais craint les sables mouvants (décapage des racines), d'où il est chassé (dunes d'Essaouira) pour céder sa place au genévrier rouge.) [5].

La figure 1 montre l'aire de répartition de l'arganier au Sud-Ouest marocain.

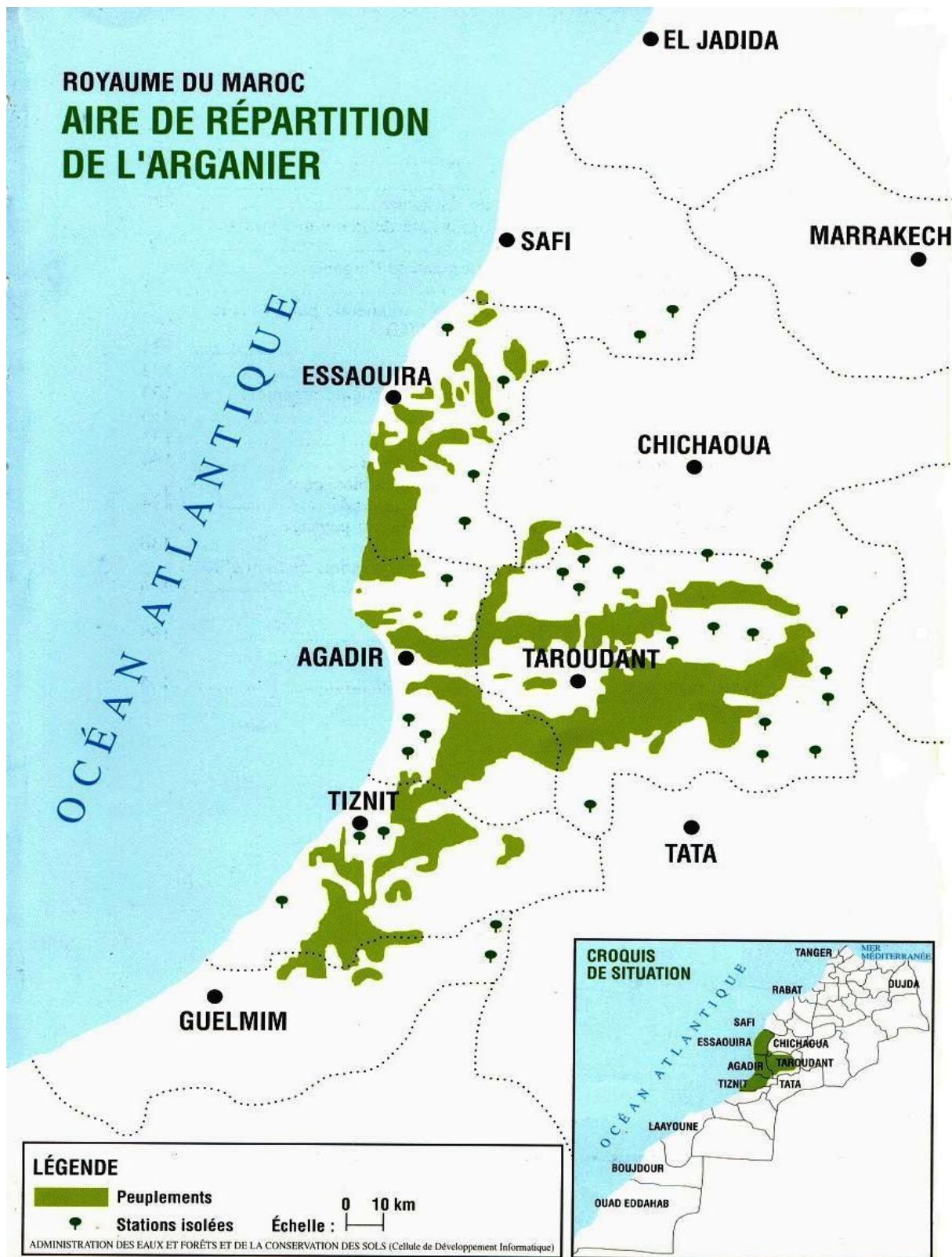


Figure 1: Carte de l'aire de répartition de l'arganier au Sud-Ouest marocain [2].

Connu pour sa rusticité [5,6], l'arganier se développe dans des conditions écologiques très variées et souvent difficiles sous forme de peuplements épars le plus souvent, constituant ainsi un véritable rempart contre l'avancée du désert. En plus de son rôle de protection des sols [7] de l'environnement et de la biodiversité [8], l'arganier joue un rôle socio-économique

important [9-11]. Selon Benzyane (1995) [12], l'arganeraie procure du bois de feu estimée à 400000 m³/an, des unités fourragères (500 UF/ha/an) et une production d'huile d'argan (3500 à 4000 tonnes/an).

La valorisation de l'huile d'argan et de ses dérivés apparaît être le moyen de relancer un développement durable pour une foresterie rurale intégrée, pouvant améliorer les revenus des ayants droits en les rendant plus responsables de la préservation de l'arganier.

Sa production est restée traditionnelle malgré l'émergence des circuits commerciaux d'approvisionnement pour répondre à la demande croissante de la part d'une clientèle régionale et nationale.

Cette huile avait fait l'objet de beaucoup d'études dans le passé [13-19]. Mais ce n'est que durant les années 98 que cette filière a connu un essor important grâce à la découverte des vertus alimentaires, cosmétiques et thérapeutiques de cette huile [17, 20-23].

Il s'est avéré que cette huile a des propriétés protectrices, adoucissantes, cicatrisantes et nettoyantes pour la peau, les cheveux et les ongles, et aussi des propriétés d'antivieillesse prématurée des cellules. De même, elle est aussi utilisée en médecine traditionnelle et qu'elle a des effets anti-inflammatoires et analgésiques [24]. L'acide linoléique qu'elle contient est déterminant dans la prévention de l'infarctus du myocarde [10] et dans la réduction du taux de cholestérol [25]. Aussi, sa teneur en polyphénols contribuerait à la prévention du cancer de la prostate [26].

L'essor important de développement qu'a connu l'huile d'argan s'est traduit par l'incitation à la création de coopératives féminines [27].

2. Caractéristiques physico-chimiques et usage de l'huile d'argan et l'arganier

2.1 Fonctions et usages de l'arganier :

L'arganier est un arbre multi-usages, chaque partie ou production de l'arbre est utilisable et représente une source de revenu pour les usagers. En plus de son rôle économique, l'arganier joue un rôle écologique important pour lutter contre la désertification.

a. Rôle économique de l'arganier :

Il convient de rappeler que les forêts d'arganier qui s'étendent sur une grande surface dont les conditions naturelles d'implantation sont dures. Pourtant, cet arbre offre divers produits d'un intérêt économique certain.

• Le bois

Le bois de l'arganier est blanc jaunâtre et dur, il constitue un combustible de haute qualité dont le rendement en charbon dépasse un quintal par m³ de bois coupé. Ce bois réputé de qualité a été utilisé dès le début du siècle, pour l'approvisionnement des grandes

agglomérations (Casablanca, Marrakech). Il est aussi une source pour les constructions rurales familiales (poutres, perchettes). Cette utilisation abusive du bois de l'arganier a été à l'origine d'une disparition de grandes étendues de l'arganeraie.

• L'exploitation du fruit

La principale richesse de l'arbre est son fruit, c'est une baie sessile formée d'un péricarpe charnu (pulpe) et d'un pseudo-endocarpe (noyau), ou sont incluses les graines. Selon la forme et la dimension, six types de fruits ont été distingués : fusiforme, ovale, ovale apicule, goutte, arrondi, globuleux.

La production fruitière varie selon l'âge de l'arbre, la densité du peuplement, le milieu et la pluviométrie. L'ordre de grandeur est de 8 kg/arbre/an, soit 1 312 000 tonnes/an pour l'arganeraie avec une densité moyenne de 200 souches/ha, [2].

La récolte du fruit se déroule entre mi-mai et mi-août, des systèmes de protection de la production fruitière contre le parcour, sont mis en place par les populations. Les fruits ainsi collectés sont exposés au soleil en couche mince afin de sécher la pulpe. L'huile d'argan est le produit principal obtenu du fruit, les autres sous-produits sont la pulpe, la coque, et les tourteaux.

La pulpe

C'est le péricarpe charnu de la noix d'argan, de couleur jaune brun clair et elle tourne au brun foncé au cours de sa dessiccation après récolte. Elle est destinée essentiellement à l'alimentation du bétail et représente de 55% à 75% du poids frais du fruit.

En vue de déterminer la valeur fourragère de la pulpe de l'argan des analyses ont été réalisées par Sandret [28], la composition chimique de la pulpe est résumée dans le tableau 1

Tableau 1: Composition chimique de la pulpe (% m/m) d'après Sandret (1956) [28]

| | |
|----------------------------------|-------------------------------|
| Humidité | 15 à 25 % suivant les régions |
| Cendres | 2 à 4 % |
| Glucides solubles réducteurs | 15 à 25 % |
| Glucides solubles non réducteurs | 5 à 15 % |
| Hémicellulose | 9 à 13 % |
| Cellulose | 6 à 12 % |
| Lipides | 5 à 10 % |
| Protides | 6 à 8 % |

La coque

La coque est un sous-produit du concassage, et représente 90 % de la noix. Elle est utilisée par les ménages riverains aux zones de l'arganier pour la combustion.

L'amendon

Il a une utilisation intéressante : on en retire l'huile qui constitue une source de revenu pour les exploitations de l'arganier. De l'extraction de l'huile résultent des résidus appelés tourteaux ; ils sont donnés au bétail comme nourriture.

Le tourteau

C'est le résidu de l'extraction, utilisé comme aliment pour bovins soumis à l'engraissement. Il est riche en glucides et protéines et renferme un important groupe pharmacodynamique constitué de saponines, [29]. La composition du tourteau est donnée au tableau 2.

Tableau 2 : Composition des tourteaux à l'état frais (% , m/m) [30]

| | |
|-------------------------|----------------|
| Humidité | 26.30 % |
| Matières grasses | 18.85 % |
| Matières azotées | 24.60 % |
| Glucides | 9.00 % |
| Cendres | 3.65 % |
| Celluloses | 17.60 % |

2.2. Caractéristiques physico-chimiques de l'huile d'argan

Les principales constantes physico-chimiques de l'huile d'argan vierge relevées dans la littérature sont rassemblées dans le tableau 3.

Tableau 3 : Principales constantes physico-chimiques de l'huile d'argan vierge.

| Constantes | Norme marocaine de l'huile d'argan [31] | H.A [32] | H.P.T [32] | H.P.N.T [32] |
|---|---|----------|------------|--------------|
| Indice d'acidité % | ≤3.3 | 0.89 | 0.75 | 0.86 |
| Indice de peroxyde meq/kg | ≤20 | 0.71 | 0.7 | 0.4 |
| Indice d'iode gI%g | - | 107.97 | 107.46 | 109.05 |
| Indice de réfraction à 20 °c | 1.463-1.472 | - | - | - |
| Indice de saponification | 189-199.1 | - | - | - |
| Insaponifiable % | ≤1.1 | - | - | - |
| Phosphore en ppm | - | 9.92 | 32.91 | 31.53 |
| Teneur en eau et matières volatiles % m/m | ≤0.2 | 0.22 | 0.089 | 0.098 |
| Teneur en impuretés insolubles %m/m dans l'éther de pétrole | ≤0.3 | 0.13 | 0.29 | 0.22 |

La composition chimique de l'huile d'argan s'est révélée intéressante par la nature de sa fraction glycérique (99%) et de son insaponifiable 1%.

La fraction glycéridique constitue 99% de l'huile totale ; les triglycérides sont majoritaires à 95%.

L'insaponifiable contient des hydrocarbures et des carotènes 37,50%, des tocophérols 7,50%, des alcools triterpéniques 20%, des méthyl-stérols et stérols 20% et des xantophylles 6,50% [18 ,33-34].

Références bibliographiques

- [1] Benzyane, M. 1988. Le rôle socio-économique et environnemental de l'arganier. In : Journées d'étude sur l'arganier, Essaouira 23-24
- [2] M'hirit O., Benzyane M., Benchekroun F., EL Yousfi S. M. et Bendaanoun M. 1998. L'arganier : une espèce fruitière-forestière à usages multiples. Mardaga. Belgique, 145 p.
- [3] El Aich, Bourbouze, A., & El Aïch, A. 2005. L'élevage caprin dans l'arganeraie: l'utilisation conflictuelle d'un espace. Cahiers Agricultures, 14, 447-453.
- [4] Boudy, 1950. Economie forestière Nord-africaine. Monographies et traitements des essences forestières. Tome 11, fasc., Edit, Larose Paris.
- [5] Emberger L., 1925. Le domaine naturel de l'arganier. Bull. soc. Bot. Fr. 72.770-774
- [6] Boudy P., 1952. Guide forestier en Afrique du Nord, Maison Rustique, Paris , 185-194.
- [7] M'hirit O., 1989. Formation forestière continue : Thème « l'arganier », Division de Recherche et d'Expérimentations Forestières, 13-17 Mars 1989, Rabat, 32-56.
- [8] Radi N., 2003. L'arganier : arbre du sud-ouest Marocain, en péril, à protéger. Thèse d'état, novembre 2003. Faculté de pharmacie, université de Nantes.
- [9] Benchekroun F. & Buttoud G., 1989. L'arganeraie dans l'économie rurale du Sud-Ouest marocain. Forêt Méditerranéenne, 21-23.
- [10] Cherki M., Berrougui H., Drissi A., Adlouni A., Khalil A., 2006. "Argan oil: Which benefits on cardiovascular diseases?" Pharmacological Research 54(1), 1-5.
- [11] Nouaïm R., 2005. L'arganier au Maroc, entre mythes et réalités. Paris : Le Harmattan, 2005.
- [12] Benzyane M., 1995. Rôle socio-économique et environnemental de l'arganier in Actes de journées d'étude sur l'arganier. Essaouira du 29 au 30 septembre Maroc.
- [13] Jaccard, 1926 .L'aganier ,sapotacée oléagineuse du Maroc, Pharmaceutica Acta helvetiae, 11, 203-209
- [14] Berrada, 1972. Contribution à l'étude chimique de l'huile d'argania spinosa .Thèse 3ème cycle. Université Ibnou Zohr Agadir
- [15] Collier A., Lemaire B., 1974. Etude des caroténoïdes de l'huile d'argan. Cah. Nutr. Diét. 9(4),300-301.
- [16] Er-Rafia M., 1975. L'arganeraie et ses problèmes. Thème personnel, ENGREF Nancy, 38 p + Annexes 19 p.
- [17] Rojas L., Quideau S., Pardon P., Charrouf Z., 2005. Colorimetric evaluation of phenolic content and GC-MS characterization of phenolic composition of alimentary and cosmetic argan oil and press cake. J. Agric. Food Chem., 2005, 53, 9122-9127.
- [18] Farines M., Soulier M., Charrouf Z., 1984. Etude de l'huile des graines d'Argania spinosa

(L.); Sapotaceae. I. La fraction glycéridique. Rev. Franç. Des Corps Gras. 31 (7/8):283-286.

[19] Zarrouk K., Smoughen S., Maurin R., 1987. Etude de la pulpe du fruit de l'arganier (*Argania spinosa*) du Maroc. Matière grasse et latex. Actes Ins. Agro. Vet. Rabat 7,17-22.

[20] Belcadi, 1994. Etude des variations du système antioxydant cellulaire en fonction de l'âge et de l'apport alimentaire d'acides gras polyinsaturés, chez le rat. Influence particulière de l'ingestion de l'huile d'argan.

[21] Khallouki F., Younos C., Soulima ni R., Ost er T., Charrouf Z., Spiegelha Ider B., Bartsch H., 2003. Consumption of argan oil (Morocco) with its unique profile of fatty acids, tocopherols, squalene, sterols and phenolic compounds should confer valuable cancer chemopreventive effects. Eur J Cancer Prev 12:67.

[22] Rahmani M., 2005. Composition chimique de l'huile d'argane « vierge ». Cahiers d'Etudes et de Recherches Francophones / Agricultures. Volume 14, Numéro 5, 461-5.Menendez.

[23] Charrouf Z., Guillaume D., 2008. Argan oil: Occurrence, composition and impact on human health. European Journal of Lipid Science and Technology. Volume 110, Issue 7, pages 632–636, No.7

[24] Alaoui, K. Lagorge, JF Cherrah, Y. Hassar, M. Amarouch H. et Roquebert. J. 1998 Analgesic and anti-inflammatory activity of *Argania spinosa* saponins. Ann. Pharm. Fr., .56, 220-8.

[25] Berrougui H., Ettaib A., Herrera G., Alvarez D., Hmamouchi M., 2003. Hypolipidemic and hypocholesterolemic effect of argan oil (*Argania spinosa* L.) in *Meriones shawi* rats. J Ethnopharmacol. 89,15-18.

[26] Bennani H., Fiet J., Adlouni A., 2009. Impact de l'huile d'argan sur le cancer de la prostate : étude de l'effet antiprolifératif des polyphénols. Revue Francophone des Laboratoires , Numéro 416S1, 23-26.

[27] Bouzemouri B., 2007. Problématique de la conservation et du développement de l'arganier. In : Charrouf Z. (synthèse coord. par), l'arganier levier du développement humain du milieu humain du milieu rural marocain. Actes du colloque international, Rabat, 27-28 Avril, 15-19.

[28] Sandret, 1956.La pulpe d'argan, composition chimique et valeur fourragère. Ann, Rech, Forest, Maroc, Tome 4, fasc.1, 153-177, Rabat

[29] Cotton, 1988. Etude sur la noix d'argan; nouveau principe immédiat;J. Pharm. Chim., 18, 298.

[30] Charrouf Z, 1991.Contribution à l'étude chimique de l'huile d'argania spinosa (L.) skeels. Phytochem. Rev., 1, 345- 354

[31] Charrouf, Z. El kabous, A. Nouaim, R. Bensouda Y. et Yaméogo. R. 1997 Etude de la composition chimique de l'huile d'argane en fonction de son mode d'extraction. Al Biruniya Revue marocaine de Pharmacognosie. 13, 35- 39.

[32] Norme marocaine homologuée de corps gras d'origines animale et végétale, huiles d'argane N M 08.5.090. Ministère de l'Industrie, du Commerce, de l'Energie et des Mines 2002.

[33] Charrouf. M. 1984. Contribution à l'étude chimique de l'huile d'*Argania spinosa* (L.) (Sapotaceae). Thèse Sciences Univ. de Perpignan.

[34] Charrouf, Z. Guillaume, D. et Driouch, A. 2002 l'arganier, un atout pour le Maroc. *Biofutur*.220,54-57.

B.L'HUILE D'OLIVE

1 .Généralités

L'olivier couvre une superficie d'environ 620.000 ha (220 ha en irrigués, 150.000 ha en bour et 250.000 ha en montagnes) [1] représentant plus de la moitié de la superficie arboricole nationale, et constituant, de ce fait, la principale culture fruitière du pays. Il est présent à travers l'ensemble du territoire national en raison de ses capacités d'adaptation à tous les étages bioclimatiques, allant des zones de montagne aux zones arides et sahariennes. Il assure, de ce fait, des fonctions multiples de lutte contre l'érosion, de valorisation des terres agricoles et de fixation des populations dans les zones marginales.

La production des olives est de 630.000 tonnes dont une partie est destinée à la production d'huile d'olive soit 80.000 tonnes [2] provenant de la trituration de 400.000 tonnes d'olives par les unités semi-modernes et modernes (systèmes de super-presses et de centrifugation à trois phases), de 20.000 tonnes d'olives par les unités à deux phases équipées par des centrifugeuses horizontales et de 70.000 tonnes d'olives par les maâsras, l'autre partie est utilisée pour la préparation des olives de table, soit 110.000 tonnes d'olives [2].

La forte demande en huile d'olive vierge de bonne qualité est due non seulement à ses vertus de santé mais également à ses propriétés organoleptiques [3].

a .Production

L'huile d'olive a été cultivée depuis des milliers d'années dans les pays méditerranéens, et a joué un rôle important dans le régime alimentaire, l'économie et la culture de la population dans cette région [4-6].

La répartition de la production mondiale de l'huile d'olive en 2011

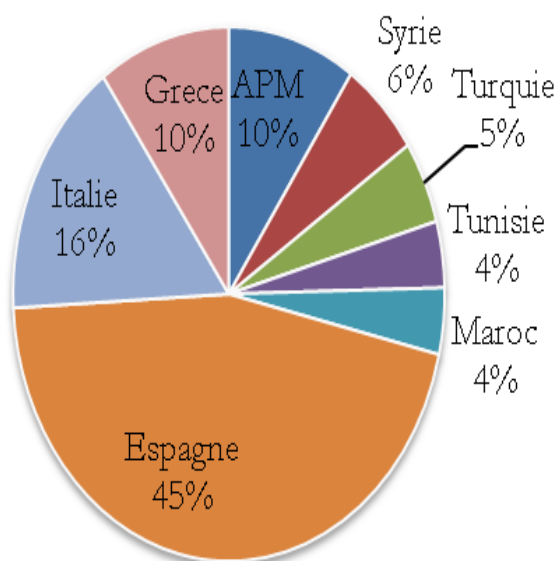


Figure 2 : Evolution de la production mondiale de l'huile d'olive en 2011 (**World Oil 2011**). Les estimations du Conseil Oléicole International (COI) pour la campagne 2011-2012 indiquent une production mondiale autour de 3,1 millions de tonnes.

b. Consommation

En ce qui concerne la consommation, les conclusions suivantes peuvent être tirées [7]. Des données par l'IOOC:

- 1- 90 % de la production mondiale d'huile d'olive sont consommés dans les pays producteurs eux-mêmes (Figure 3).
- 2- les pays producteurs-exportateurs principaux ont une consommation plutôt régulière qui se situe entre 1.520.000 et 1.580.000 tonnes.
- 3- les pays producteurs-importateurs montrent une tendance vers une consommation croissante.
- 4- les pays uniquement importateurs montrent également une légère tendance vers l'augmentation de la consommation d'huile d'olive.

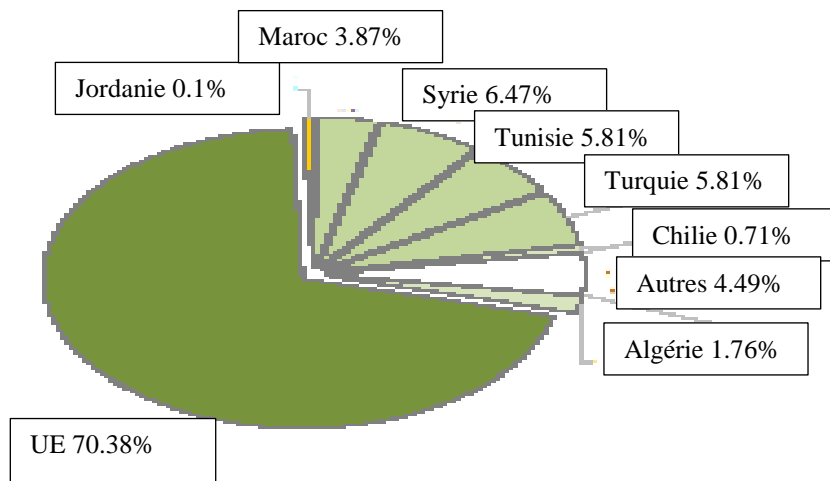


Figure 3 : Principaux pays consommateurs en 2011/2012 en tonnes - source COI - graph. A.Paris/Afidol)

À première vue, les huiles d'olive semblent donc avoir une importance limitée en terme de statistiques mondiales. En comparaison avec d'autres huiles végétales comestibles, les huiles d'olive représentent 3 % de la consommation humaine et sont largement dépassées par l'huile de soja (27 %), l'huile de palme (20 %), l'huile de graine de colza (15 %), et l'huile de tournesol (12 %). De même, dans le commerce international, les huiles d'olive ne représentent pas plus que 2 % du volume d'huiles végétales comestibles vendues [8].

Cependant, l'importance économique des huiles d'olive se développe considérablement dans certaines régions du monde. Ce secteur représente 25 % du revenu total des fermiers des pays méditerranéens. Ces pays sont responsables de 74,4 % de la production mondiale, l'Espagne et l'Italie étant les producteurs principaux (Figure 2).

2 .Composition chimique des huiles végétales et facteurs influant la qualité de l'huile

2.1. Facteurs influençant la qualité de l'huile d'olive

La qualité de l'huile d'olive est influencée par plusieurs facteurs climatiques, géographiques, pédologiques et génétiques ; ainsi que par le mode d'extraction, les pratiques culturales et les conditions de stockage [9].

- **Facteurs climatiques**

De nombreuses recherches ont été faites pour en garantir la pureté, l'authenticité et la qualité de l'huile. Kler, [10] a observé en premier l'influence de la variété, de la région de culture et des conditions climatiques sur la composition et les caractéristiques.

Le climat a une influence importante sur la maturité, donc sur la composition chimique de l'huile d'olive. Des travaux de recherche réalisés par Aparicio ont prouvé que les structures cycliques de l'huile d'olive sont affectées par le climat [11].

La lumière et la température affectent la concentration en acide gras de l'huile d'olive. La composition en acides gras insaturés et principalement en acide linoléique augmente avec la diminution de la température [12].

La température et les précipitations ont un effet sur le comportement physiologique de l'olivier, et par conséquent sur les caractéristiques chimiques de l'huile. [11; 13 ; 14].

- **Facteurs géographiques**

Les olives cultivées dans différentes zones géographiques présentent des caractéristiques différentes.

La qualité de l'huile d'olive est affectée par l'altitude. Cette dernière affecte la composition de l'huile d'olive en acides gras, principalement l'acide oléique. Les olives cultivées à haute altitude donnent des huiles riches en acide gras mono insaturés, bien que les olives cultivées à faible altitude donnent des huiles riches en acides gras saturés donc plus stables. De même elle présente un effet sur l'acidité, l'indice de peroxyde, l'indice d'iode et la teneur en polyphénols [10-12 ; 15-17].

- **Facteurs pédologiques**

L'influence du sol sur la qualité de l'huile d'olive est un phénomène assez complexe. Plusieurs facteurs tels que la nature du sol, le pH et la composition chimique se mettent en jeu et peuvent influencer la qualité d'une huile.

En général, les huiles provenant des sols calcaires ont une acidité plus basse que celles des sols argileux [9].

- **Le mode d'extraction**

Il existe de très nombreuses données sur la caractérisation et la composition en acides gras de l'huile d'olive d'origines variées (Bulgarie, Espagne, France, Grèce, Italie, Maroc, Portugal et Tunisie). Plusieurs études ont été réalisées pour corréler la composition chimique de l'huile d'olive, sa qualité avec l'origine géographique ainsi qu'avec les procédés d'extraction [18-28]. Les procédés d'extraction connus peuvent altérer la qualité de l'huile d'olive en affectant sa stabilité durant la conservation. Deux systèmes d'extraction sont utilisés : le système traditionnel ou procédé discontinu et les systèmes modernes ou procédés continus.

Les huiles produites par le système continu contiennent des taux élevés en polyphénols, ce qui leur confère une résistance contre l'oxydation pour un stockage à long terme. Les huiles d'olives extraites par le système continu ont une acidité inférieure à 1%. Tandis qu'en mode discontinu, elle devient légèrement supérieure à 1% [12].

- **Les pratiques culturales**

L'huile extraite des olives dépend du stade de maturation auquel les fruits ont été récoltés. La composition en acides gras saturés comme l'acide palmitique décroît durant la période de maturation, tandis que celle de l'acide oléique et linoléique augmente [29].

La date et le type de récolte, le système d'irrigation, la lutte contre les ravageurs et le type de cultivar doivent aussi être pris en considération dans l'étude de l'effet des pratiques culturales sur la qualité de l'huile.

- **Les conditions de stockage**

L'huile d'olive doit être conservée soigneusement à tous les stades jusqu'au moment où elle est mise à la consommation.

Les conditions de stockage (la durée, la température,...) ont un effet sur l'acidité, l'indice de peroxyde, la stabilité, la couleur, la composition en acide gras et en tocophérols [30-32].

Pour la classification des huiles d'olive, des méthodes chimiométriques ont été appliquées [33-34]. Parmi ces méthodes, l'Analyse en Composante Principale (ACP), l'Analyse Discriminante, et la Classification Hiérarchique ont été utilisées.

Composition de l'olive

L'olive est une drupe. Sa composition dépend de sa variété, du sol et du climat. Le contenu de l'olive est composé du noyau (17,3%-23% massique), d'amandon (2%-5,5% massique), d'épicarpe (2%-2,5% massique) et de la pulpe (71,5%-80,5% massique) [35] figure 4.

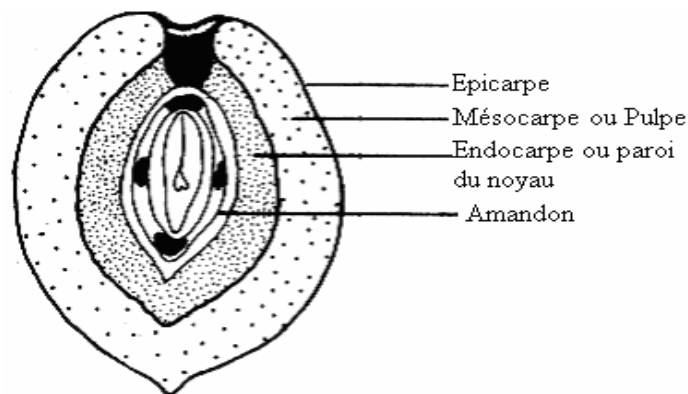


Figure 4. Composition de l'olive [35]

Composition des huiles

La composition des huiles alimentaires d'origine végétale est dominée par les triglycérides (95 – 98 %) avec différentes compositions en acides gras. Les composantes minoritaires (2 –5 %) appartiennent à des groupes chimiques diverses, tels que alcools, esters, hydrocarbures, composés phénoliques, tocophérols, tocotriénols, pigments, phospholipides, entre autres.

Les acides gras des triglycérides les plus fréquents dans les huiles sont :

- polyinsaturées: linoléique (C18 :2) et linoléique (C18 :3)
- mono-insaturées: oléique (C18 :1)
- saturées: stéarique (C18 :0) et palmitique (C16 :0)

Cette variabilité est responsable des différentes caractéristiques physico-chimiques des huiles, notamment la résistance à l'oxydation.

La consommation de certaines huiles avec un contenu élevé en acides polyinsaturés est bénéfique dû à son influence sur l'équilibre hormonal, le renforcement du système immunitaire et la prévention du cancer, du diabète, de l'obésité, de l'arthrite, entre autres. Plusieurs études [36-40] ont révélé les effets positifs sur la santé humaine d'une alimentation riche en acides gras insaturés (acide oléique en particulier),

En revanche, la consommation des huiles peut être très mauvaise pour la santé quand elles ont été dégradées, soit naturellement lors du stockage dans le temps, soit suite à des procédés industriels ou culinaires qui peuvent accélérer des transformations néfastes.

2.2. Composition chimique des huiles végétales

Le tableau 4 montre la composition en acides gras de quelques huiles végétales.

Les principaux triacylglycérides de l'huile d'olive sont POO (18,4 %), SOO (5,1 %), POL (5,9 %), OOO (43,5 %), OOL (6,8 %) (P: acide palmitique; O: acide oléique; S: acide stéarique; L: acide linoléique) [41].

Tableau 4 : Pourcentage des acides gras dans quelques huiles végétales [8].

| Acides Gras | | Huile d'olive | Huile de colza | Huile de noix de coco | Huile de maïs | Huile de coton | Huile de tournesol | Huile d'argan [42] |
|-----------------------|--------|---------------|----------------|-----------------------|---------------|----------------|--------------------|--------------------|
| Acide myristique | C14:0 | <0,05 | 0,1-0,2 | 16,5-20,8 | 0-0,3 | 0,6-1,0 | 0-0,1 | |
| Acide palmitique | C16 :0 | 7,5-20,0 | 3,0-5,0 | 8,2-10,2 | 9,1-16,8 | 21,0-26,8 | 5,5-7,7 | 11,5-15 |
| Acide palmitoléique | C16 :1 | 0,3-3,5 | 0,2-0,6 | | 0-0,3 | 0-1,3 | 0-0,3 | |
| Acide heptadécanoïque | C17 :0 | < 0,3 | | | | | | |
| Acide heptadécénoïque | C17 :1 | < 0,3 | | | | | | |
| Acide stéarique | C18 :0 | 0,5-5,0 | 1,0-2,0 | 2,3-3,4 | 1,4-3,0 | 2,0-3,3 | 2,8-6,5 | 4,3-7,4 |
| Acide oléique | C18 :1 | 55,0-83,0 | 52,0-67,0 | 4,3-8,1 | 20,0-38,0 | 14,0-22,0 | 14,0-38,0 | 43-49,1 |
| Acide linoléique | C18 :2 | 3,5-21,0 | 16,0-24,8 | 0,7-2,0 | 39,5-65,0 | 46,5-58,0 | 48,2-74,2 | 28-36 |
| Acide linoléinique | C18 :3 | 0,9 | 6,5-14,0 | tr | 0,6-1,4 | 0-0,4 | 0-0,1 | <0,4 |
| Acide arachidique | C20 :0 | 0,6 | 0,2-0,8 | 0,1 | 0,3-0,7 | 0,2-0,5 | 0,2-0,4 | 0,3-0,5 |
| Acide eicosénoïque | C20 :1 | C20:1Z0,4 | 0,9-2,4 | 0-tr | 0,2-0,4 | 0-0,1 | 0-0,2 | |
| Acide béhénique | C22 :0 | <0,2 | 0,1-0,5 | | 0-0,5 | 0-0,6 | 0,7-1,3 | |
| Acide lignocérique | C24 :0 | <0,2 | 0-0,2 | | 0-0,3 | | 0-0,4 | |
| Note: tr=trace | | | | | | | | |

Références bibliographiques

- [1] MADRPM, 2006. Données générales sur l'agriculture marocaine. Les cultures arboricoles. Direction de la Statistique. Ministère de l'Agriculture et de la Pêche Maritime
- [2] Le Conseil Oléicole International COI/OT/NC n° 1 Décembre 2004
- [3] Luaces, P. Pérez, A. G., et Sanz, C. 2003. Role of olive seed in the biogenesis of virgin olive oil aroma. *Journal of Agricultural and Food Chemistry*, 51, 4741-4745.
- [4] Keys A. 1980. Seven countries: A multivariate analysis in death and coronary heart disease, Harvard Univ. Press, Cambridge, Massachusetts, USA,
- [5] Viola P. 1987. Olive oil and health, International Olive Oil Council, Madrid, Spain.
- [6] Muriana F. J. G., Ruiz-Gutierrez V. 1989. Effects of w-6 and w-3 polyunsaturated fatty acids ingestion on rat liver membrane-associated enzymes and fluidity. *J. Nutr. Biochem.* 3, 659-663.
- [7] Boskou D. 1996. Olive Oil – Chemistry and Technology, AOCS PRESS, Champaign, Illinois., 2-83, 101-133.
- [8] Harwood John., Ramon. Aparicio. 2000. Handbook of olive oil – Analysis and properties, An Aspen publication, Aspen Publishers, Inc., Gaithersburg, Maryland, 1-513.
- [9] Çavusoglu, A. Otkar, A. 1994. Les effets des facteurs agronomiques et des conditions de stockage avant la mouture sur la qualité de l'huile d'olive. *Olivæ, Science et technique*, n° 52, 18-24.
- [10] Kler, J. 1992. «Histoire des corps gras» In «Manuel des corps gras» KARLESKIND.A. Tome 1, Edition : Lavoisier, Paris, 10-11.
- [11] Aparicio, R. Ferreiro, L. Alonso, V. 1994. Effect of climate on the chemical composition of virgin oil, *Analytica Chimica Acta.*, 292,235-241,
- [12] Mouawad, M. 2005. Effet de l'altitude, sol et pressoir sur la qualité d'huile d'olive, mémoire de master en biochimie, Université Saint Esprit de Kaslik,.1, 7-14, 23-28, Liban.,
- [13] Pannelli, G. Servili, M. Selvaggini, R. Baldioli, M. et Montedoro, G.F. 1994. Effect of agronomic and seasonal factors on olive (*Olea europaea* L.) production and on the qualitative characteristics of the oil. *Acta Horti*, 356, 239-243.
- [14] Ryan, D. Robards, K. et Lavee, S. 1998. Evolution de la qualité de l'huile d'olive. *Olivæ*, 72, 23-41.
- [15] Ollivier, D. 2003. Recherche d'adultération dans les huiles végétales –application à la qualité des huiles vierges et notamment de l'huile d'olive. *OCL*, 10,4, 315-320,

- [16] Ouaini, N. Medawar, S. Daoud, R. Ouaini, R. Chebib, H. Rutledge, D. Estephan. N. 2005. Etat actuel des huileries d'olive au Liban – Potentiel de production, *New Medit journal*, Pub. 4/2005, Liban.
- [17] Ranalli, A. De Mattia, G. Patumi, M. Proietti, P. 1999. Quality of virgin oil as influenced by origin area, *Grasa y aceites*, 50,4, 249-259.
- [18] Aparicio, R. Albi, T. Lanzon, A. et Navas, M.A. 1987. Sexia, un sistema experto para la identificación de aceites : base de datos de zonas olivareras. *Grasas y Aceites*, 38,1, 9-14.
- [19] Leardi, R. et Paganuzzi, V.1987. Caratterizzazione dell'origine di oli di oliva extravergini mediante metodi chemiometrici applicati alla frazione sterolica. *Rivista Italiana dell Sostanza Grasse*, LXIV, 131-136.
- [20] Fiorina, M. et Nizzi, F. 1991. The spread of olive farming. *Olivae*, 44, 9.
- [21] Ferreiro, L. et Aparicio, R. 1992. Influencia de la altitud en la composición química de los aceites de oliva vírgenes de Andalucía. Ecuacione/8*s matemáticas de clasificación. *Grasas y Aceites*, 43,3, 149-156.
- [22] Gigliotti, C. D'Agghetta, A. et Sidoli, A. 1993. Indagine conoscitiva sul contenuto trigliceridico di oli extra vergini di oliva di varia provenienza. *Rivista Italiana dell Sostanza Grasse*, LXX, 483-489.
- [23] Tsimidou, M. Papadopoulos, G. et Boskou, D.1992. Determination of phenolic compounds in virgin olive oil by reverse-phase HPLC with emphasis on UV detection. *Food Chemistry*, 44, 53-60.
- [24] Ulberth, F. et Buchgraber, M. 2000. Authenticity of fats and oils. *Eur. J. Lipid. Sci. Technol*, 102, 687-694.
- [25] Di Giovacchino, L. Sestili, S. et Di Vincenzo, D. 2002. Influence of olive processing on virgin olive oil quality. *Eur. J. Lipid Sci. Technol.*, 104,587–601
- [26] Dhifi, W., Angerosa, F., Serraiocco, A., Oumar, I. Hamrouni, I. et Marzouk, B. 2005. Virgin olive oil aroma: Characterization of some Tunisian cultivars. *Food Chemistry*, 93, 697-701.
- [27] Galeano Diaz, T. Durán Merás, I. Sanchez Casas, J. et Alexandre Franco, M.F. 2005. Characterization of virgin olive oils according to its triglycerides and sterols composition by chemometric methods. *Food Control*, 16, 339-347.
- [28] Pardo, J.E. Cuesta, M.A. et Alvarruiz, A. 2007. Evaluation of potential and real quality of virgin olive oil from the designation of origin “Aceite Campo de Montiel” (Ciudad Real, Spain). *Food Chemistry*, 100, 977–984.

- [29] Cimato, A. 1990. La qualité de l'huile d'olive et les facteurs agronomiques, *Olivae*, 31, 20-31.
- [30] Pereira, J. A. Casal, S. Bento, A. Oliveira, M. B. P. P. 2002. Influence of olive storage period on oil quality of three portuguese cultivars of *Olea europea*, Cabrançosa, Madural, and Verdeal Transmontana, *Food Chemistry*, 50, 6335-6340.
- [31] Sacchi, R. Mannina, L. Fiordiponti, P. Barone, P. Paolillo, L. et Patumi, M. 1998. Characterization of Italian extravirgin olive oils using H-NMR spectroscopy. *Journal of Agricultural and Food Chemistry*, 46, 3947-3951.
- [32] Ranalli, A. Modesti, G. Patumi, M. et Fontanazza, G. 2000. The compositional quality and sensory properties of virgin olive oil from a new olive cultivar-I-77. *Food Chemistry*, 69, 37-46.
- [33] Ollivier D., Boubault E., Pinatel C., Souillol S., Guérère M., Artaud J. 2004. Analyse de la fraction phénolique des huiles d'olive vierges. *J. Annales des falsifications, de l'expertise chimique et toxicologique*, 965, 169 - 196.
- [34] Runcio, A. Sorgona, L. Mincione, A. Santacaterina, S. et Poiana, M. 2008. Volatile compounds of virgin olive oil obtained from Italian cultivars grown in Calabria. Effect of processing methods, cultivar, stone removal, and antracnose attack. *Food Chemistry*, 106, 735-740
- [35] Nefzaoui A. 1984. Importance de la production oléicole et des sous-produits de l'olivier. In : *Etude de l'utilisation des sous-produits de l'olivier en alimentation animale en Tunisie. Étude FAO production et santé animales* 43, Rome.
- [36] Berry E.M. et Rivlin R.S., 1997. Dietary fatty acid in the management of diabetes mellitus. *American Journal of Clinical Nutrition*, 66, 4, 991-997.
- [37] Jing M., Folsom A.R., Lewis L., Eckfeldt J.H., et Ma J., 1997. Relation of plasma phospholipid and cholesterol ester fatty acid composition of artery intima-media thickness: the arterosclerosis risk in communities (ARIC) study. *American Journal of Clinical Nutrition*, 65, 2, 551-559.
- [38] Kinter M., Spitz D.R., et Roberts R.J., 1996. Oleic acid incorporation protect cultured hamster fibroblast from oxygen-induced cytotoxicity. *Journal of Nutrition*, 126, 12, 2952-2959.
- [39] Krajcovicova-kudlakova M., Simonsic R., Bederova A., et Klvanova J., 1997. Plasma fatty acid profile and alternative nutrition. *Annal of Nutrition and Metabolism*, 41, 6, 365-370.

[40] Po Huang C., Nomura A.M.W., Stemmermann G.N., et Chyou P.H., 1996. A perspective study of colon and rectal cancer among Hawaii Japanese men. *Annals of Epidemiology*, 6, 4, 276-282.

[41] Fedeli, E. et Camurati, F., 1981. Valorisation des margines et des grignons épuisés par récupération de quelques composants In : Séminaire International sur la Valorisation des Sous Produits de l'olivier. PNUDFAO. Monastir, Tunisie, Décembre 1981, 111-113.

[42] Khallouki, F, Younos C, Soulimani R, et al. Consumption of argan oil (Morocco) with its unique profile of fatty acids, tocopherols, squalene, sterols and phenolic compounds should confer valuable cancer chemopreventive effects. *Eur J Cancer Prev* 2003;12:67-75.

CHAPITRE II : DETERMINATION DES ACIDES GRAS DES HUILES PAR CHROMATOGRAPHIE EN PHASE GASEUSE

1. Réaction de transestérification des huiles

1.1. Introduction

La transestérification des triglycérides n'est pas un processus nouveau. Elle date de 1853 lorsque Patrick et Duffy ont mené cette réaction beaucoup d'années avant que le premier moteur Diesel soit fonctionnel [1].

Cette réaction a été l'objet de recherches intensives grâce aux utilisations diverses de ces produits comprenant la synthèse des polyesters dans l'industrie des polymères [2], la synthèse des intermédiaires pour l'industrie pharmaceutique [3], le durcissement des résines dans l'industrie de peinture et dans la production du biodiesel en tant qu'alternatif du diesel.

Les esters d'acides gras peuvent être préparés à partir d'huiles végétales, contenant majoritairement des triglycérides, selon deux voies de production : l'estérification d'acides gras obtenus par hydrolyse des triglycérides ou la transestérification directe de l'huile végétale, qui est la voie de synthèse la plus largement employée dans l'industrie [4]. Les matières premières utilisées dans les deux cas sont principalement des huiles végétales de tournesol, de colza, de soja ou de palme. Pour notre étude bibliographique, nous nous intéresserons plus particulièrement à la réaction de transestérification des triglycérides avec un alcool. Cette réaction, schématisée sur la figure 5, conduit à la formation d'esters alkylés d'acides gras et de glycérol.

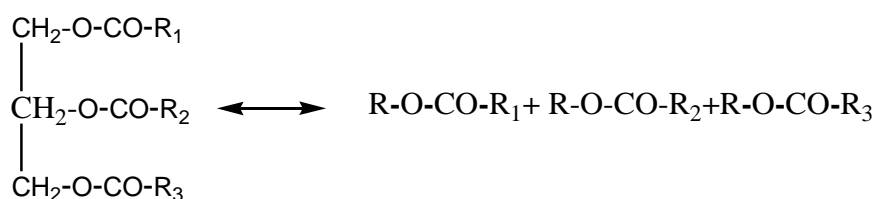


Figure 5 : Réaction générale de transestérification de triglycérides avec un alcool (R-OH) avec R₁, R₂ et R₃ des groupements alkyles présents dans les acides gras

Nommée aussi alcoolyse, c'est la réaction entre un ester et un alcool conduisant à un ester différent [5].

Elle se décompose en trois réactions successives et réversibles. La première étape est la conversion des triglycérides (TG) en diglycérides (DG) ; suivie de la conversion des diglycérides en monoglycérides (MG), qui sont finalement transformés en glycérol (GL). A chaque étape, une molécule d'ester d'acide gras est formée. (Figure 6) :

- 1ère étape : $TG + R-OH \leftrightarrow DG + R1-COO-R$
 2ème étape : $DG + R-OH \leftrightarrow MG + R2-COO-R$
 3ème étape : $MG + R-OH \leftrightarrow GL + R3-COO-R$

Figure 6 : Les trois étapes successives et réversibles de la réaction de transestérification

Chromatographie en phase gazeuse: constituée de 3 parties

- L'injecteur (1) : L'échantillon est introduit avec une microsiringue (de 1 à 10 µL), à travers un septum en élastomère, dans la chambre de vaporisation. L'injecteur a une double fonction : il porte l'échantillon à l'état de vapeur, puis il l'amène dans le flux gazeux en tête de colonne.
- La colonne : elle est placée dans une enceinte à température régulée (2). Elle se présente sous la forme d'un tube de silice, enroulée sur lui-même et de longueur allant de 1 à plus de 60 m.
- Entraînés par un gaz vecteur inerte, les analytes étudiés sont séparés en fonction de leur capacité d'interaction avec la phase stationnaire.
- Le détecteur (3) : il s'agit du module qui va permettre de détecter voire d'identifier les composés en sortie de colonne, relié à un intégrateur ou un ordinateur (4) sur lequel apparaît le chromatogramme.

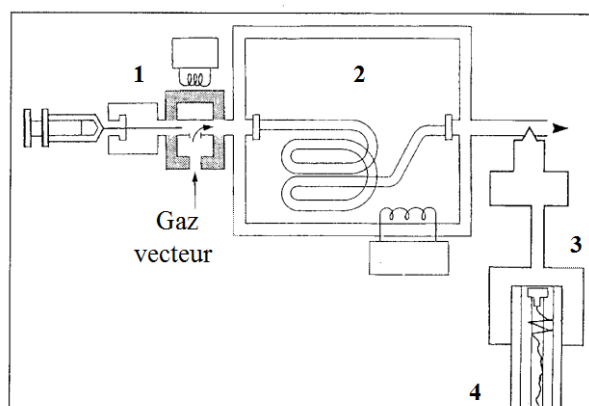


Schéma d'appareillage

1.2. Les aspects thermodynamiques

La transestérification des huiles végétales est une réaction qui se déroule en milieu biphasique : les triglycérides et l'alcool ne sont pas totalement miscibles. Suivant la température à laquelle est conduite la réaction de transestérification, l'apparition dès les premiers instants d'ester alkylique peut conduire à une homogénéisation du milieu réactionnel [6]. Celui-ci est alors constitué d'une seule phase jusqu'à ce que l'avancement de la réaction soit suffisant pour entraîner la formation de deux nouvelles phases ; L'une contenant

majoritairement de l'ester d'acide gras et l'autre du glycérol. Plusieurs auteurs ont travaillé sur les diagrammes d'équilibre de phase [7].

2. Les différentes techniques d'analyse d'un mélange complexe

2.1 La Chromatographie en Phase Gazeuse (CPG)

La chromatographie introduite en 1952 est une technique incontournable pour l'analyse des gaz. Elle permet la séparation des constituants d'un mélange et d'effectuer la plupart des dosages quelle que soit la teneur, grâce au couplage avec divers détecteurs.

La CPG est une méthode d'analyse par séparation qui s'applique aux composés gazeux ou susceptibles d'être vaporisés par chauffage sans décomposition [8]. C'est la technique de séparation la plus utilisée dans le domaine des huiles végétales, car elle permet d'effectuer l'identification des constituants à partir d'échantillons de l'ordre du milligramme voire du microgramme. Les progrès technologiques réalisés dans le domaine des colonnes capillaires, des phases stationnaires et des détecteurs (FID) ont contribué à rendre la CPG incontournable pour l'analyse des huiles végétales.

Cependant, une reproductibilité parfaite des indices de rétention est difficile à obtenir et ne peut être observée que sur des chromatogrammes réalisés sur une période courte avec des conditions expérimentales rigoureusement identiques. Les variations les plus importantes sont observées lorsqu'on compare les indices de rétention obtenus au laboratoire avec ceux de la littérature, particulièrement pour ce qui concerne la colonne polaire [9].

De plus, les nombreux cas de co-élution montrent que la comparaison des indices de rétention et même la méthode des surcharges ne suffisent pas à identifier formellement un composé. Le développement des phases stationnaires chirales (colonne de cyclodextrines) et de la CPG multidimensionnelle a permis de surmonter certaines difficultés rencontrées dans la séparation et l'identification des composés dans les huiles essentielles [9,10]. Ainsi, la CPG bidimensionnelle (CPG/CPG), mettant en ligne deux colonnes capillaires, permet la séparation, l'identification et la quantification de composés minoritaires pouvant coéluer avec des composés plus abondants. L'échantillon est injecté dans une pré-colonne, puis les composés qui coeluent sont transférés dans une deuxième colonne pour être séparés. L'efficacité de cette méthode est illustrée par Marriott et coll. [9] qui ont appliqué la CPG bidimensionnelle à l'étude d'une huile essentielle de vétiver (*Vetiveria zizanioides*).

2.2. Application de la chromatographie en phase gazeuse

La détection de l'adultération de l'huile d'olive par CPG est réalisée en comparant la composition de leur acide gras [11] et l'analyse chimométrique des acides gras [12]. L'analyse chimométrique des triglycérides et les acides gras des olives françaises est employé pour la détermination de la variété et l'origine géographiques [13].

L'utilisation de colonnes chirales en CPG bidimensionnelle permet l'analyse énantiomérique des constituants d'un mélange complexe. Dans de nombreux cas, la détermination de la distribution énantiomérique est utilisée pour la détection d'adultérations dans les aromes [14,15]. Enfin, la GPG/CPG chirale a été utilisée par König et coll. [16] pour identifier le (-)-d-amorphène dans une huile essentielle de vétiver.

Les techniques analytiques basées sur la chromatographie en phase gazeuse (CPG) permettent l'analyse d'extraits de plantes [17]. La chromatographie en phase gazeuse couplée à la spectrométrie de masse (CPG/MS) a permis d'améliorer la sensibilité et l'identification des composés, comme dans le cas d'études sur des feuilles d'*Arabidopsis thaliana* [18].

Les techniques de couplage de la CPG

Les techniques de couplage de la CPG avec des techniques spectroscopiques (SM, IRTF) ont permis une approche beaucoup plus précise de l'identification.

Chaque constituant est caractérisé par deux indices de rétention polaire et apolaire, calculés à partir d'une gamme d'alcanes ou plus rarement d'esters méthyliques linéaires, à température constante (indice de Kováts) [19] ou en programmation de température (indices de rétention) [20]. Les temps de rétention, bien que spécifiques d'un composé, ont tendance à varier d'une analyse à l'autre, notamment du fait du vieillissement des colonnes.

Ils sont ensuite comparés avec ceux de produits de référence (mesurés au laboratoire ou décrits dans la littérature). Toutefois, il est fréquent d'observer des variations, parfois importantes, lorsque l'on compare les indices de rétention obtenus au laboratoire et ceux de la littérature (en particulier sur colonne polaire).

Malgré tout, ceci ne peut suffire à une bonne identification, sans l'apport du couplage entre la CPG et une technique d'identification spectroscopique : en général la spectrométrie de masse (CPG/SM) ou plus rarement l'IRTF. La combinaison de ces deux techniques complémentaires, est applicable à l'analyse d'un grand nombre de substances organiques, aussi bien gazeuses que liquides.

Les techniques de couplage de la CPG avec des techniques spectroscopiques (SM, IRTF) ont permis une approche beaucoup plus précise de l'identification.

Le couplage CPG/SM.

Le couplage CPG/SM a permis un gain en sensibilité et une grande sensibilité du signal par rapport aux couplages précédemment développées avec d'autres systèmes de détections [21].

Ce couplage est relativement plus facile puisque les composantes de l'échantillon sont élues sous forme de vapeur à la sortie de la colonne CPG puis ensuite directement introduite dans le vide du spectromètre de masse. Il offre de multiples possibilités analytiques et combine les avantages des deux techniques : le pouvoir séparatif de la CPG et le pouvoir d'identification de la SM. L'un de ces avantages est son pouvoir de confirmation basé sur une bibliothèque regroupant les spectres d'ionisations obtenus par impact électronique pour un très grand nombre de composés [22]

Ainsi le couplage CPG/SM en mode impact électronique (SM-IE) est la technique la plus utilisée dans le domaine des huiles essentielles. Il permet de connaître, dans la grande majorité des cas, la masse moléculaire d'un composé et d'obtenir des informations structurales relatives à une molécule à partir de sa fragmentation. Dans la source d'ionisation les molécules sont bombardées à l'aide d'électrons, conduisant ainsi à la formation des ions en phase gazeuse. Les ions sont ensuite dirigés vers la partie analytique de l'appareil.

Comment utiliser au mieux cette séparation ?

Si on ne sépare pas les différents composants, on s'aperçoit que l'ionisation du produit entier va fournir un résultat complexe impossible à analyser, contenant des informations relatives aux différents composés.

Après la séparation chromatographique, on a tout intérêt à analyser composé par composé. Grâce au gaz vecteur, et suivant le temps de rétention des différents composés, ceux-ci vont passer un par un dans un spectromètre de masse où ils vont subir une ionisation électronique (énergie de l'ordre de 70eV) et chimique (on brise la molécule avec une énergie plus grande de l'ordre de 90 eV). Le spectre obtenu permettra de définir un graphe donnant l'importance des différents fragments en fonction de leur rapport m/z (rapport masse sur la charge). L'étude du spectre à l'aide d'une banque de données permet de définir le composé en jeu et son importance dans le produit final. Cette méthode est très sensible et est capable d'identifier la plupart des composés présents malgré leur gamme de concentration qui peut être très étendue. De plus cette méthode sensible, est spécifique et nécessite un minimum de prélèvement de

l'échantillon : ceci se révèle très utile lorsque l'on veut faire des analyses dans le domaine de l'archéologie où il est très souvent délicat d'utiliser un gros échantillon.

□ **Mode de détection : la spectrométrie de masse est constituée de 3 parties**

- Analyseur quadripolaire : L'étude des trajectoires suivies, dans une enceinte où règne le vide, permet de déterminer le rapport masse sur charge (m/z) de l'ion. En balayant les amplitudes U et V, la trajectoire des ions peut-être stabilisée ou déstabilisée en fonction de leur rapport m/z . Seuls les ions qui ont une trajectoire stable vont pouvoir traverser le quadripôle et arrivent au détecteur, les autres étant éjectés.

-Chambre d'ionisation : en sortie de colonne, l'intégralité des analytes entrent dans la chambre d'ionisation. Le mode d'ionisation est l'impact électronique (EI)



Ces ions formés dans la source sont ensuite accélérés et focalisés par une électrode chargée positivement (le repousseur) vers l'analyseur : le quadripôle.

- Détecteur : Le détecteur recueille alors ces ions séparés par l'analyseur en fonction de leur rapport m/z . Puis un micro-ordinateur va assurer le traitement des données et fournir un spectre de masse. L'EI est un procédé reproductible qui permet des comparaisons spectrales à l'aide de bibliothèques de spectres (**NIST**).

Références Bibliographiques

- [1] Demirbas, A., Springer-Verlag, 2008. Biodiesel : a realistic fuel alternative for dieselengines
- [2] Meyer, U., Hoelderich, W.F., 1999. Transesterification of methyl benzoate and dimethyl terephthalate with ethylene glycol over basic zeolites Appl. Catal. A.,178, 159-166.
- [3] Barrault, J., Pouilloux, Y., Clacens, J.M., Vanhove, C., Bancquart, S., Catalysis and fine chemistry, Catalysis Today, 2002. 75, 177-181.
- [4] Gervajio G.C., 2005. Fatty Acids and Derivatives from Coconut Oil. In: Bailey's Industrial Oil and Fat Products, Ed. Wiley.
- [5] Allinger, N.L., Cava, M. P., De Jongh, D. C., Johnson, C. R., Lebel, N. A., Stevens, C.L., 6ème ed, 1983. 2,553.
- [6] Liu X.J., Piao X.L., Wang Y.J. and Zhu S.L., 2008. Liquid-Liquid Equilibrium for Systems of (Fatty Acid Ethyl Esters + Ethanol + Soybean Oil and Fatty Acid Ethyl Esters + Ethanol + Glycerol). Journal of Chemical and Engineering Data; 53(2), 359-362.
- [7] Lanza M., Neto W.B., Batista E., Poppi R.J. and Meirelles A.J.A., 2008. Liquid- Liquid Equilibrium Data for Reactional Systems of Ethanolysis at 298.3 K. Journal of Chemical & Engineering Data; 53(1), 5-15.
- [8] Grundschober F., 1991. The identification of individual components in flavouring and flavoured foods, Z. Lebensm. Unters. Forsh., 192, 530-534.
- [9] Marriott P., Shellie R., Fergus J., Ong R., Morrison P., 2000. High resolution essential oil analysis by using comprehensive gas chromatographic methodology, Flavour Fragr. J., 15, 225-239.
- [10] Marriott P., Shellie R., Cornwell C., 2001. Gas chromatographic technologies for the analysis of essential oils, J. Chromatogr. A, 936, 1-22.
- [11] Cercaci, L., Rodriguez-Estrada, M. T., & Lercker, G. 2003. Solidphase extraction-thin-layer chromatography-gas chromatographymethod for the detection of hazelnut oil in olive oils by determination of esterified sterols. Journal of Chromatography A, 985, 211–220.
- [12] Dourtoglou, V. G., Dourtoglou, T., Antonopoulos, A., Stefanou, E., Lalas, S., & Poulos, C. 2003. Detection of olive oil adulteration using principal component analysis applied on total and regio FA content. Journal of the American Oil Chemists Society, 80, 203–208.
- [13] Ollivier, D., Artaud, J., Pinatel, C., Durbec, J. P., & Guerere, M. 2003. Triacylglycerol and fatty acid compositions of French French virgin olive oils. Characterization by chemometrics.51(19),5723-31.

- [14] Casabianca H., Graff J.B., Jame P., Perruchietti C., 1995. Application of hyphenated techniques to the chromatographic authentication of flavors in food products and perfumes, *J. High Resol. Chromatogr.*, 18, 279-286.
- [15] Casabianca H., Graff J.B., Faugier V., Fleig F., Grenier C., 1998. Enantiomeric distribution studies of linalool and linalyl acetate. A powerful tool for authenticity control of essential oils, *J. High Resol. Chromatogr.*, 21, 107-112.
- [16] König W.A., Bülow N., Saritas Y., 1999. Identification of sesquiterpene hydrocarbons by gas phase analytical methods, *Flavour Fragr. J.*, 14, 367-378.
- [17] Roessner U, Luedemann A, Brust D, Fiehn O, Willmitzer TLL, Fernie AR 2001. Metabolic profiling allows comprehensive phenotyping of genetically or environmentally modified plant systems. *The Plant Cell* 13, 11-29.
- [18] Fiehn O, Kopka J, Dörman P, Altmann T, Trethewey RN, Willmitzer L 2000. Metabolite profiling for plant functional genomics. *Nature Biotechnology* 18, 1157-1161.
- [19] Kováts E., 1965 Gas Chromatographic characterization of organic substances in the retention index system, in *Advances in Chromatography*, Chap. 7, 229-247.
- [20] Van Den Dool H., Kratz P.D., 1963, A generalization of the retention index system including linear temperature programmed gas-liquid partition chromatography, *J.Chromatogr.*, 11, 463-471.
- [21] Hercegovà A., Domotorovà M., Matisovà E.,Kirchner M.,Otrekal R.,Stefuca 2005 Fast gas chromatography with solid phase extraction clean-up for ultratrace analysis of pesticide residues in baby food. V.,*J.Chrom. A*, 1084 46-53
- [22] Hermàndez F., Pozo O.J.,Sancho J.V.,Bijlsma L.,Barreda M .,Pitarch E 2006.Multi-residue liquid chromatography tandem mass spectrometry determination of 52 non gas chromatography amenable pesticides and metabolites in different food commodities ,*J.Chrom.A*, 1109,242-252

Chapitre 3: LA SPECTROSCOPIE INFRAROUGE ET SES APPLICATIONS DANS L'ANALYSE DES HUILES

1. Généralité

Les problèmes d'authenticité dans la filière agroalimentaire ne sont pas récents et remontent même aux civilisations grecque et romaine.

La grande différence de prix entre les huiles vierges et les huiles végétales peut entraîner des tentations de falsification consistant à introduire tout ou partie d'huile de moindre coût dans une huile d'un prix supérieur. Ce type de fraude peut constituer une menace majeure pour la santé des consommateurs (plus de 400 décès et la maladie de 20.000 personnes en Espagne en 1990) sont les syndromes d'huile toxique [1].

Cette opération frauduleuse a été étudiée en combinant la spectroscopie IR à des outils chimométriques.

L'huile d'olive pure a pu être distinguée d'huile d'olive adultérée par ajout d'autres huiles végétales : tournesol ou noisette [2].

Contrairement à des techniques uni-analytiques comme par exemple la pH-métrie où à chaque mesure on obtient un seul point ou résultat, une méthode multi-analytique peut acquérir de nombreuses données simultanément.

2. La spectroscopie infrarouge

2.1 Le rayonnement infrarouge

Les techniques de spectroscopie sont rapides, d'un coût de fonctionnement faible et peuvent apporter de nombreuses informations [3 -4].

Le rayonnement infrarouge (IR) fut découvert en 1800 par Frédéric Wilhelm Herschel. Ces radiations localisées au-delà des longueurs d'onde dans le rouge, sont situées entre la région du spectre visible (400 à 900 nm) et des ondes hertziennes (10^{10} à 10^{11} nm). Le domaine infrarouge s'étend de 0,7 μm à 50 μm . Il est classé en 3 grandes catégories, le proche infrarouge (0,7 à 2,5 μm), le moyen infrarouge (2,5 à 25 μm) et le lointain infrarouge (25 à 50 μm). La figure 7 représente le spectre électromagnétique qui s'étend des ondes radio aux rayons cosmiques.

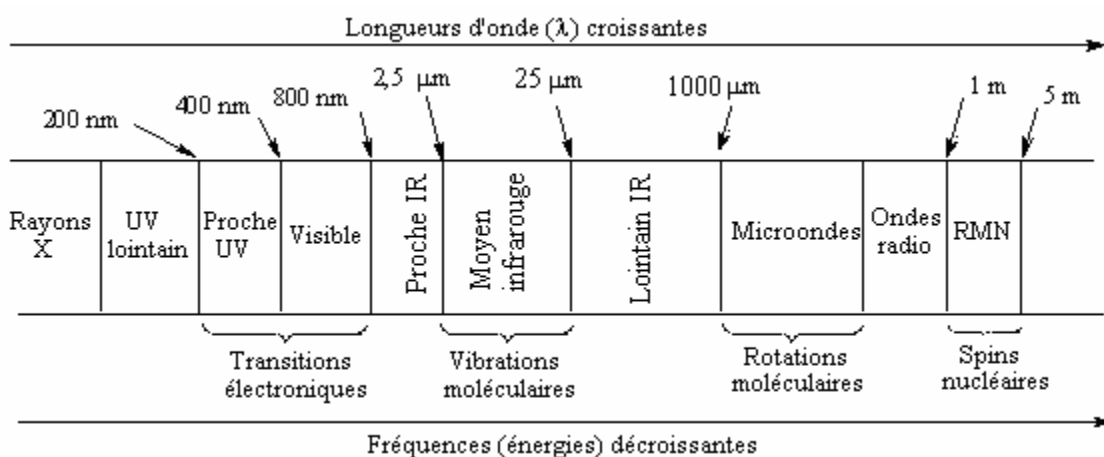


Figure 7 : Le spectre électromagnétique [5]

Le principe de la spectroscopie infrarouge a été détaillé notamment par Bertrand et par Rouessac [6 ; 7]

De nombreuses propriétés du rayonnement électromagnétique peuvent être décrites par la théorie des ondes classiques qui utilise des paramètres tels que la longueur d'onde, la fréquence, la vitesse et l'amplitude. Toute onde électromagnétique est caractérisée par deux paramètres, sa fréquence et sa longueur d'onde, qui sont proportionnelles à l'énergie E.

L'énergie d'une radiation est liée à une longueur d'onde ou à une fréquence par la relation.

$$E = h\nu = hc / \lambda \quad \text{Eq. 1}$$

Où **h** est la constante de Planck ($h=6,626 \cdot 10^{-34}$ J·s), **c** la vitesse de la lumière ($3 \cdot 10^8$ m·s⁻¹), **v** la fréquence et **λ** la longueur d'onde exprimée en mètres.

Tableau 5 : longueurs d'onde (μm) et nombres d'onde (cm⁻¹) pour les régions du proche, moyen et lointain infrarouge.

| | longueurs d'onde (μm) | nombres d'onde (cm ⁻¹) |
|---------------------|-----------------------|------------------------------------|
| proche infrarouge | 0,7-2,5 μm | 13300-4000 cm ⁻¹ |
| moyen infrarouge | 2,5-25 μm | 4000-400 cm ⁻¹ |
| lointain infrarouge | 25-1000 μm | 400-10 cm ⁻¹ |

Dès 1924, l'énergie du rayonnement moyen infrarouge est la plus utilisée pour l'analyse puisque l'ensemble des molécules possèdent les fréquences d'absorption caractéristiques et les vibrations moléculaires primaires dans cette gamme et elle permet de garder une sorte d'empreinte digitale. La Spectroscopie infrarouge moyen est basée sur l'étude de l'interaction du rayonnement infrarouge avec la matière. Ainsi, la relation entre l'absorption d'un rayonnement moyen-IR par une molécule et sa structure moléculaire sont

mises en évidence. Dans le PIR on trouve essentiellement les bandes harmoniques des vibrations fondamentales rencontrées dans le moyen infrarouge. Même si les régions du proche IR et de l'IR lointain ont suscité un certain intérêt, l'utilisation de la spectroscopie moyen-IR reste la plus adaptée pour l'élucidation de la composition moléculaire d'un composé.

2.2 Le spectromètre Infrarouge à Transformée de fourrier

Un spectromètre IRTF comporte essentiellement cinq parties (figure 8) [8- 9] :

- Une source lumineuse : par exemple, le Globar est utilisé pour le moyen IR. Il est constitué d'un bâtonnet en carbure de silicium (SiC) et fonctionne à des températures de l'ordre de 1500°C. Il rayonne un spectre continu entre 10000 et 250 cm^{-1} .
- Un dispositif permettant de générer les interférences: l'interféromètre de Michelson
- Un compartiment échantillon qui permet d'accueillir plusieurs types d'accessoires (porte-échantillon) dépendant de la technique spectroscopique employée et le mode de mesures utilisé (réflexion ou transmission).
- Un détecteur ou capteur photosensible: le spectromètre IRTF peut comporter un ou plusieurs détecteurs comme les détecteurs DTGS (Deuterated Tri-glycine Sulfate) ou photoélectrique (générant une différence de potentiel par l'absorption de photons) comme les détecteurs MCT (Mercure Cadmium Tellure) qui sont constitués d'un monocristal en alliage de mercure cadmium tellure déposé sur un support inerte.
- Enfin, le convertisseur analogique numérique qui interroge le détecteur à des intervalles réguliers et transforme le signal analogique en un signal numérique

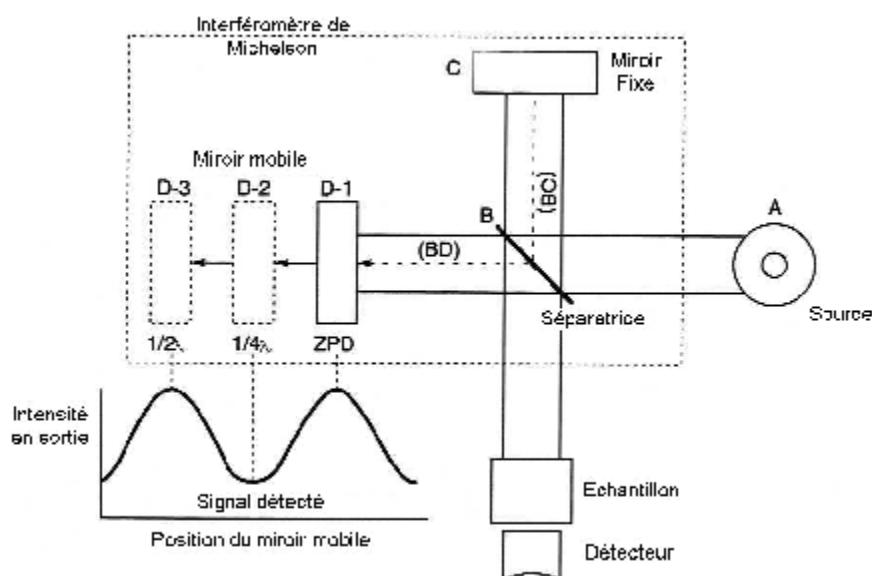


Figure 8 : Schéma de principe d'un spectromètre infrarouge à transformée de Fourier [8]

a. Fonctionnement du spectromètre IRTF

Pour résumer, dans tous les spectromètres IRTF, le faisceau infrarouge provenant de la source est dirigé vers l'interféromètre qui va moduler chaque longueur d'onde du faisceau à une fréquence différente. Dans l'interféromètre le faisceau lumineux arrive sur la séparatrice. La moitié du faisceau est alors dirigée sur le miroir fixe, le reste passe à travers la séparatrice et est dirigé sur le miroir mobile. Quand les deux faisceaux se recombinent, des interférences destructives ou constructives apparaissent en fonction de la position du miroir mobile. Le faisceau modulé est alors réfléchi des deux miroirs vers l'échantillon, où des absorptions interviennent. Le faisceau arrive ensuite sur le détecteur pour être transformé en un signal électrique.

L'interférogramme est la somme de toutes les fréquences du faisceau. Cet interférogramme est ensuite converti en un spectre infrarouge par l'opération mathématique de transformée de Fourier.

b. Réflexion totale atténuée (ATR)

Le deuxième mode de mesure en spectroscopie IRTF est la réflexion totale atténuée, communément appelée ATR (attenuated total reflection). La spectrométrie infrarouge est une méthode d'analyse dont le principe repose sur l'absorption de la lumière infrarouge par la matière organique. L'étude des spectres permet de tirer deux types d'informations :

- *informations qualitatives* : les liaisons chimiques absorbent à une longueur d'onde caractéristique. Des tables IR permettent d'attribuer les absorptions aux différents groupes chimiques en présence et donc d'identifier la ou les molécules présentes.
- *informations quantitatives* : l'intensité de la bande d'absorption rend compte de la concentration du groupe chimique caractéristique de cette absorption (loi de Beer-Lambert).

Le principe de l'ATR est détaillé sur la figure 9. Ce principe a été détaillé par Harrick [10].

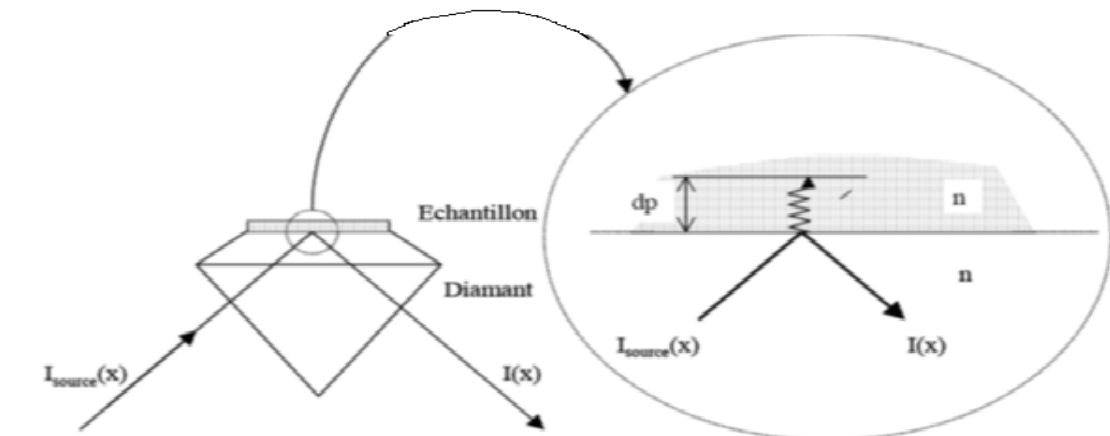


Figure 9 : Principe de la réflexion totale atténuée (ATR) sur diamant

L'échantillon est placé sur un cristal transparent en infrarouge, ce cristal est de type ZnSe, diamant, germanium mais d'indice de réfraction élevé appelé élément à réflexion interne. $I_{source}(x)$ est l'intensité globale émise par la source en fonction du retard et $I(x)$ est l'intensité après réflexion sur l'échantillon. En première approximation, d'après la loi de Descartes, le faisceau IR initial traverse le cristal (diamant dans la figure 9) et subit une réflexion totale à l'interface cristal-échantillon puis est dirigé vers le détecteur [8].

En effet, la particularité du cristal, outre sa haute résistance, est qu'il possède un indice de réfraction élevé ($n_2=2,43$ pour le diamant par exemple) et dans la plupart des cas supérieur à celui de l'échantillon (n_1). En réalité, le phénomène peut être perturbé par l'existence d'une onde progressive appelée évanescente. Celle-ci pénètre de quelques micromètres dans l'échantillon se trouvant en contact direct avec le cristal. De ce fait, une partie de l'énergie lumineuse est retenue et la réflexion totale est dite atténuée [8].

L'ATR est la technique permettant l'étude des échantillons complexes, voire impossibles à analyser par les méthodes spectroscopiques conventionnelles: comme les échantillons opaques en raison de leur forte absorption (solution aqueuses, cristaux liquides visqueux, fibres textiles) ou de leur épaisseur. Le phénomène de réflexion totale atténuée permet donc d'examiner uniquement la partie superficielle de l'échantillon étudié.

3. La Spectroscopie Moyen infrarouge (IR) et ses applications

Applications de spectromètres FT-IR sont énormes dans presque tous les domaines, tels que l'industrie [9-13], la recherche [14-17], et la science médicale [18-20]. La technique est un outil rapide et rentable pour la surveillance de routine de plusieurs constituants tel que le saccharose, le glucose, le fructose, l'acide citrique, etc. dans la production de jus courant par exemple [21,22]

En spectroscopie infrarouge, on s'intéresse aux différentes vibrations des liaisons. Elle est appliquée pour des composés organiques et inorganiques soit à l'état gazeux, liquide ou solide.

La spectroscopie MIR est caractérisée par une résolution fine des pics qui lui confère plusieurs avantages, elle permet de garder une sorte d'empreinte digitale. Tout d'abord, la finesse des raies permet d'identifier facilement le type de liaison rencontré et même de quantifier un composé en fonction de l'intensité des signaux. En plus de fournir des informations qualitatives sur les groupes fonctionnels, les spectres IR peuvent fournir des informations quantitatives, comme la concentration. Un autre avantage important des spectres MIR est qu'ils apportent des informations sur l'état physicochimique de différents composants comme la cristallinité, l'oxydation, ou encore la phase, qui peuvent être utiles pour étudier les changements causés par le traitement ou le stockage [23]. Le coût d'analyse très faible par échantillon [3-4]. Ces méthodes nécessitent une préparation simple des échantillons et elles peuvent être prédictives [24-26].

Le développement des applications MIR en agroalimentaire est récent, suite à l'apparition d'appareils basés sur la transformée de Fourier et de nouveaux dispositifs de présentation des échantillons plus pratiques [27].

3.1. Application de la spectroscopie Moyen Infrarouge

Plusieurs problèmes d'authentification ont été étudiés par ce type de méthode ; par exemple la quantification des adultérants de l'huile d'olive [28], ou encore la discrimination des différents types de fruits dans les purées [29] ou les confitures [30], et enfin la discrimination des différentes variétés de café [31,32]. Au niveau du contrôle qualité de la production, une étude polonaise [33] s'est intéressée à l'évolution de la teneur en phospholipides lors de la production d'huile de colza. L'application des techniques spectroscopiques IR et RAMAN combinées avec les méthodes chimiométriques est une nouvelle approche pour la détermination de l'authenticité de l'huile d'olive. Obeidat, et al [34] ont étudié l'adultération d'huile d'olive avec de l'huile de tournesol et de l'huile de maïs. Fernando et al, [35] ont détecté l'adultération de l'huile d'olive avec de l'huile de noisette par le couplage spectrométrie de masse et les techniques de régression multi variée telle que la PLS et l'ACP.

Dernièrement, divers procédés ont été proposés pour établir l'authenticité des huiles végétales et à détecter le niveau de falsifications comme la Spectrophotométrie UV sur la

base de l'absorbance dans les régions de $208 \pm 310 \pm 210$ et 320 nm a été utilisée par Passaloglou-Emmanouillidou [36] pour détecter la falsification de l'huile d'olive. La Spectrométrie IR a été utilisée pendant de nombreuses années dans l'huile et l'industrie des graisses [37, 38,39]. Cependant, ces méthodes présentent des limites en tant que procédé de détection spectrale parce que les différences entre la plupart des huiles végétales qui contiennent les mêmes acides gras, principalement C16 ou C18 et leur TG (C50, C52, C54), sont assez petites.

La spectroscopie IRTF a également été utilisée pour distinguer les différentes origines géographiques de l'huile d'olive extra vierge [40] et différentes variétés génétiques [41].

3.2 Application de la spectroscopie moyen infrarouge à l'étude des huiles.

Etude du degré d'insaturation des lipides

La spectroscopie MIR est devenue une méthode de routine officielle pour doser les acides gras "*cis*" et "*trans*" dans les huiles et les matières grasses [42].

Ahlers S K D. *et al.*, [43] ont étudié le degré d'insaturation des acides gras libres purs et estérifiés. En utilisant l'ATR. Belton P. S. *et al.*, (1988) [44] ont dosé les acides gras "*trans*" insaturés dans les huiles et les matières grasses.

Dans la pratique, le degré d'insaturation des lipides est apprécié par le dosage de l'indice d'iode. Cet indice, basé sur une réaction d'addition, peut être utilisé pour déterminer quantitativement l'insaturation globale des chaînes grasses [45]. Anderson B. A. *et al.*, (1974)[46] ont développé une méthode spectroscopique pour déterminer avec précision le degré d'insaturation des lipides dans les beurres et les huiles.

Etude de la longueur de chaîne des lipides

Bannon C. D. *et al.*, (1984) [47] ont étudié les spectres des acides gras à l'état solide dans la région $1300-1100$ cm^{-1} . Dans cette région, le nombre de bandes associées aux groupes CH_2 croît avec la longueur de la chaîne. Parkash S. *et al.*, (1975) [48] ont montré que la longueur de la chaîne d'un acide gras libre peut être caractérisée par la variation de l'intensité de la bande à 720 cm^{-1} .

3.3. Attribution des bandes d'absorption des lipides dans le moyen infrarouge

Le tableau 6 donne les attributions chimiques des principales bandes d'absorption des lipides dans le MIR.

Le Tableau 6: Les attributions chimiques des principales bandes d'absorption des lipides dans le MIR [49].

| Nombre d'onde cm^{-1} | Nature de vibration |
|--------------------------------|---|
| 3035 | élongation C=C-H " <i>trans</i> " [50] |
| 3018 | élongation asymétrique C=C-H " <i>cis</i> " [50] |
| 3007 | élongation symétrique C=C-H " <i>cis</i> " [51] |
| 2961 | élongation asymétrique CH ₃ [51] |
| 2935 | élongation asymétrique CH ₂ [50] |
| 2880 | élongation symétrique CH ₃ [50] |
| 2863 | élongation symétrique CH ₂ [50] |
| 1761 | élongation C=O monomère (acide) [52] |
| 1754 | élongation C=O ester [50] |
| 1721-1701 | élongation C=O dimère (acide) [52] |
| 1463 | déformation ("scissoring") CH ₂ [50] |
| 1381 | déformation symétrique CH ₃ [50] |
| 1190-1183 | élongation C-O acide [52] |
| 1150 | élongation C-O ester [50] |
| 1123 | déformation dans le plan CH ₃ ("rocking") [50] |
| 967 | déformation C=C-H " <i>trans</i> " [44] |
| 935 | déformation dans le plan O-H acide [44] |
| 727 | déformation CH ₂ ("rocking") [50] |

4. Application de la spectroscopie proche infrarouge

Les techniques spectroscopiques en proche infrarouge (PIR ou Near Infrared NIR) et Raman se développent de plus en plus pour la recherche de contrefaçons de médicaments tels que Combiron (sulfate de fer),

Norris était le premier à montrer les mérites de la régression multilinéaire pour extraire l'information à partir des spectres PIR, afin de déterminer la composition d'un échantillon [53]. Christy et al. [54] ont utilisé la spectroscopie proche infrarouge dans l'objectif de la quantification et la détection de l'adultération du l'huile d'olive par soja, noisette et tournesol.

Comme le seul exemple de la spectroscopie proche infrarouge, Özdemir et al, [55] ont porté une étude sur la quantification de falsification binaire et tertiaire de l'huile d'olive avec

l'huile de tournesol et l'huile de maïs avec l'inverse génétique moindres carrés partielles . La variation était entre 1,42% et 6,38% (vol / vol) pour les mélanges ternaires l'huile d'olive, de tournesol et l'huile de maïs. Dans les dernières années, l'application de la spectrométrie proche infrarouge (NIR) aux huiles et graisses est devenue plus populaire pour l'étude de la qualité et de composition. Dans les huiles d'olive, la spectroscopie NIR a été utilisée pour (i) la prévision d'identification de falsification [56], (ii) différenciation-classification des huiles végétales [57] (iii) la détermination de la valeur de l'acidité et de l'amertume [58]. L'indice d'iode des huiles est déterminé par proche infrarouge [56].

Les méthodes spectroscopiques sont largement utilisées pour l'analyse qualitative et quantitative des constituants des produits agroalimentaires [60,61]. En général, quand les signaux sont bien définis et non superposés, on peut à partir de l'intensité du signal calculer une concentration. Cependant, lors de l'analyse de matrices complexes, les signaux ne sont pas toujours bien séparés. Il n'est donc pas évident d'identifier et de quantifier les différents composants. Afin de rester dans l'objectif premier de l'analyse spectrale, qui est d'obtenir un résultat d'analyse rapide sur différents paramètres, il est alors nécessaire d'utiliser des méthodes d'extraction d'information rapide telles les méthodes chimiométriques.

Les techniques spectroscopiques peuvent être considérées comme des réponses d'absorption d'énergie en fonction de la fréquence du rayonnement concerné [61,62].

Références bibliographiques

- [1] JIMENO, S. A. 1982. The Spanish toxic symptoms. *Trends in Analytical Chemistry*, 1, 4–6
- [2] Fauhl, C., Reniero, F. & Guillou, C. 2000. ¹H NMR as a tool for the analysis of mixtures of virgin olive oil with oils of different botanic origin. *Magnetic Resonance in Chemistry* 38, 436-443
- [3] Birlouez-Aragon, I., Sabat, P. & Gouti, N. 2002. A new method for discriminating milk heat treatment. *International Dairy Journal* 12(1), 59-67.
- [4] Karoui, R., Kemps, B., Bamelis, F., De Ketelaere, B., Merten, K., Schoonheydt, R., Decuyper, E. & De Baerdemaeker, J. 2005a Development of a rapid method based on front-face fluorescence spectroscopy for the monitoring of egg freshness: 2-evolution of egg yolk. *Eur Food Res Technol*.
- [5] Dalibart, M. & Servant, S. 2000. *Spectroscopie dans l'infrarouge, techniques de l'ingénieur*, PA., 2845.
- [6] Bertrand, D. et Dufour, D. 2000. « La spectroscopie et ses applications analytiques », Ed. Tec et Doc, Paris , 31-74.
- [7] Rouessac F. et Rouessac, A. 1998 « Analyse chimique : méthodes et techniques instrumentales modernes », Durod, , 4ème Ed., 147-171.
- [8] Griffiths, P. R. 1979 dans *Laboratory Methods in Infrared Spectroscopy.*, Chp–7, eds.R.G.J. Miller and B. C. Stace. Heyden.
- [9] Notingher, I., Imhof, R.E., Xiao, P. & Pascut, F.C. 2003. spectral depth profiling of arbitrary surfaces by thermal emission decay –Fourier Transform infrared spectroscopy *Appl. Spectrosc.* 57, 1494-1501
- [10] Cran, M.J. & Bigger, S.W. 2003. Quantitative Analysis of Polyethylene Blends by Fourier Transform Infrared Spectroscopy *Appl. Spectrosc.* 57, 928-932.
- [11] Markham, J.R., Best, P.E. & Solomon, P.R. 1994. Spectroscopic Method for Measuring Surface Temperature That Is Independent of Material Emissivity, Surrounding Radiation Sources, and Instrument Calibration *Appl. Spectrosc.* 48, 265-270
- [12] Iwaoka, T., Tabata, F. & Tsutsumi, S. 1994. Use of FT-IR for the Quantification of Water in Biomolecules *Appl. Spectrosc.* 48, 818-826.
- [13] Lutz, E.T.G., Luinge, H.J., Maas, J.H. van der & Agen, R. van 1994. Analysis of Carbon Black-Filled Rubber Materials by External Reflection FT-IR Spectrometry *Appl. Spectrosc.* 48, 1021-1025.

- [14] Ouyang, H., Sherman, P.J., Paschalis, E.P., Boskey, A.L. & Mendelsihn, R. 2004. Fourier Transform Infrared Microscopic Imaging: Effects of Estrogen and Estrogen Deficiency on Fracture Healing in Rat Femurs *Appl. Spectrosc.* 58, 1-9.
- [15] Thompson, S.E., Foster, N.S., Johnson, T.J., Valentine, N.B. & Amonette, J.E. 2003. Fourier Transform Infrared Characterization of Polymers *Appl. Spectrosc.* 57, 893-899.
- [16] Horn, B.A., Qiu, J., Owen, N.L. & Feist, W.C. 1994. FT-IR Studies of Weathering Effects in Western Redcedar and Southern Pine *Appl. Spectrosc.* 48, 662-668.
- [17] Chirsty, A.A., Nodland, E., Burnham, A.K., Kvalheim, O.M. & Dahl, B. 1994. "Determination of kinetic parameters for the dehydration of calcium oxalate monohydrate by diffuse reflectance FT-IR spectroscopy", *Appl. Spectrosc.* 48, 561-568.
- [18] Heberle, J. & Zscherp, C. 1996. ATR/FTIR difference spectroscopy of biological matter with micro second time resolution *Appl. Spectrosc.* 50, 588-596.
- [19] Lewis, E.N., Gorbach, A.M., Marcott, C. & Levin, I.W. 1996. high Fidelity Fourier transform infrared spectroscopic imaging of primate brain tissue *Appl. Spectrosc.* 50, 263-269.
- [20] Sowa, M.G. & Mantsch, H.H. 1994. FTIR step –scan photo acoustic phase analysis and depth profiling of calcified *Appl. Spectrosc.* 48, 316-319.
- [21] Conde-Gallardo, A., Cruz-Orea, A. & Tomas, S.A. 2004. Interference Effects in Photoacoustic and Reflectance Spectroscopies on TiO₂/Si Structures and TiO₂ Band Gap *Appl. Spectrosc.* 58, 917–921.
- [22] Irudayaraj, J. & Tewari, J. 2003. simultaneous monitoring of organic acids and sugars in fresh and processed apple juice by fourier transform infrared attenuated total reflection spectroscopy *Appl. Spectrosc.* 57, 1599-1604.
- [23] Wilson, R. H. 1994. Spectroscopic techniques for food analysis, pp. 246 (VCH, New York,
- [24] Karoui, R., Mouazen, A.M., Dufour, E., Schoonheydt, R. & De Baerdemaeker, J. 2006. Utilisation of front-face fluorescence spectroscopy for the determination of some selected chemical parameters in soft cheeses. *Le Lait* 86(2), 155-169.
- [25] Karoui, R., Nicolaÿ, B. & De Baerdemaeker, J. 2008. Monitoring the Egg Freshness During Storage Under Modified Atmosphere by Fluorescence Spectroscopy. *Food and Bioprocess Technology* 1(4), 346-356.
- [26] Woodcock, T., Fagan, C.C., O'Donnell, C.P. & Downey, G. 2008. Application of Near and Mid-Infrared Spectroscopy to Determine Cheese Quality and Authenticity. *Food and Bioprocess Technology* 1(2), 117-129.

- [27] Downey, G. 2006. in *La spectroscopie infrarouge et ses applications analytiques* (ed. Bertrand, D.), 479-504 (TEC & DOC, Paris,
- [28] Lai, Y. W., Kemsley, E. K. & Wilson, R. H. 1995. Quantitative analysis of potential adulterants of extra virgin olive oil using infrared spectroscopy. *Food Chemistry* 53, 95-98
- [29] Defernez, M., Kemsley, E. K. & Wilson, R. H. 1995. Use of infrared spectroscopy and chemometrics for the authentication of fruit purées. *J. Agric. Food Chem.* 43, 109-113
- [30] Defernez, M. & Wilson, R. H. 1995. Mid-infrared spectroscopy and chemometrics for determining the type of fruit used in jam. *J Science of. Food and Agriculture* 67, 461- 467
- [31] Briandet, R., Kemsley, E. K. & Wilson, R. H. 1996. Discrimination of arabica and robusta in instant coffee by Fourier transform infrared spectroscopy and chemometrics. *J. Agric. Food Chem.* 44, 170-174
- [32] Kemsley, E. K., Ruault, S. & Wilson, R. H. 1995. Discrimination between coffea arabica and canephora variant robusta beans using infrared spectroscopy. *Food Chem.* 54, 321-326
- [33] Szydłowska-Czerniak, A. 2007. MIR spectroscopy and partial least-squares regression for determination of phospholipids in rapeseed oils at various stages of technological process. *Food Chem.* 105, 1179-1187
- [34] Obeidat, Safwan M.; Khanfar, Mai S.; Obeidat, Wasfy M. 2009. Classification of Edible Oils and Uncovering Adulteration of Virgin Olive Oil Using FTIR with the Aid of Chemometrics. *Australian Journal of Basic & Applied Sciences* ., Vol. 3 Issue 3, 2048-2053. 6p. 1 Diagram, 1 Chart, 2 Graphs.
- [35] Fernando Peñna, Soledad C´ardenas, Mercedes Gallego, Miguel Valc´arcel 2005. Direct olive oil authentication: Detection of adulteration of olive oil with hazelnut oil by direct coupling of headspace and massspectrometry, and multivariate regression techniques *Journal of Chromatography A*, 1074, 215–221
- [36] Passaloglou-Emmanouillidou, S.A., 1990 ."A comparative study of UV spectrophotometric methods for the detection olive oil adulteration by refine oils", *Zeitschrift für Lebensmittel – Untersuchung und – Forschung*, 191, p.132-134
- [37] Lai, Y.W. Kemsley, E.K. Wilson, R.H. 1994. Potential of Fourier transform infrared spectroscopy for the authentication of vegetable oils. *Journal of Agricultural and Food Chemistry*, 42, 1154. –1159.
- [38] Chapman, D. 1965. Infrared spectroscopy of lipids *J. Am. Oil Chem. Soc.* 42, 353.

- [39] Kochlar, S.P. Rossell, J.B. 1987. Analysis of oils and fats using infra-red spectroscopy. *Int. Anal.* 1, 23-26.
- [40] Bendini, A., Cerretani, L., Carrasco-Pancorbo, A., Gómez-Caravaca, A. M., Segura-Carretero, A., Fernández-Gutiérrez, A., et al. (2007). Phenolic molecules in virgin olive oils: a survey of their sensory properties, health effects, antioxidant activity and analytical methods. An overview of the last decade. *Molecules*, 12, 1679–1719.
- [41] Gozde Gurdeniz, Figen Tokatli and Banu Ozen .2007. Differentiation of mixtures of monovarietal olive oils by mid-infrared spectroscopy and chemometrics. *European Journal of Lipid Science and Technology*, 109, 1194–1202.
- [42] International Union of Pure and Applied Chemistry, 1987. *Standard Methods for the Analysis of Oils, Fats and Derivatives*, Method 2.401, 7th edition, Oxford. .
- [43] Ahlers S.K.D., Brett R.A., McTaggart N.G. 1953. Infrared study of the cis- and trans study isomers fC18 fatty acids. *J. Appli. Chem.*, 3, 433-443.
- [44] Belton P.S., Wilson R.H., Sadeghi H., Orabchi J., Peers K.E. 1988, A rapid method of the estimation f isolated trans double bands in oils and fats using Fourier transform infrared spectroscopy combined with attenuated total reflectance. *Lebensm. Wiss. U, Technol.* 21, 153-157.
- [45] Wolff J.P. 1968, *Manuel d'analyse des corps gras*. Azoulay eds. Paris. 114-120.
- [46] Anderson B.A., Miller R., Pallansch M.J. 1974. Measuring unsaturation in milk fat and other oils by differential infrared spectroscopy. *J. Dairy Sci.* 75, 156-159.
- [47] Bannon C.D., Crake J.D., Hilliker A.E. 1984, Analysis of fatty acid methyl esters with high accuracy and reliability. IV. Fats with fatty acids containing four or more carbon atoms. *J.Am. Oil Chem. Soc.* 62, 1501-1507.
- [48] Parkash S., Blanhard J.M.V. 1975. Infrared spectra of fatty acids and some related substances. *Spectro Chim. Acta. Part A.* , 31 A, 951-957.
- [49] Hui Li, F.R. van de Voort, A.A. Ismail, J. Sedman, R. Cox, C. Simard, and H. Buijs Discrimination of Edible Oil Products and Quantitative Determination of Their Iodine Value by Fourier Transform Near-Infrared Spectroscopy

- [50] Christy, A.A. Kasemsumran, S. Ozaki. Y.Du &Y. 2004. The detection and quantification of adulteration in olive oil by near-infrared spectroscopy and chemometrics, *J. Analytical Sciences* , 20, 935-940.
- [51] Magdi M.M., Richard E., Ecdonald J.T., David J.A., Samuel W. 1990, Identification and quantification of 9-trans, trans-12-octadecadienoic acid methyl ester and related compounds in hydrogenated soybean oil and margarines by capillary gas chromatography/matrix isolation/Fourier infrared spectroscopy. *J. Agric. Food Chem.* 38, 86-92.
- [52] Safar M. 1995. Comparaison des plages spectrales de l'infrarouge proche et moyen pour l'étude des produits agro-alimentaires. Thèse de doctorat: Spécialité: physico-chimie des bioproduits. Université de Nantes, Faculté des sciences et des techniques.
- [53] Norris K. H., Williams P. C. 1984b, Optimization of mathematical treatments of raw near-infrared signal in the measurement of protein in hard red spring wheat: Part I. Influence of particle size. *Cereal Food World.* 22, 461.
- [54] McClure W. F. 1994. Near Infrared Spectroscopy, Spectroscopic techniques for Food Analysis, ed R.H.Wilson, VCH Publishers. 13-57.
- [55] Ozdemir, D., & Ozturk, B. 2007. Near infrared spectroscopic determination of olive oil adulteration with sunflower and corn oil. *Journal of Food and Drug Analysis*, 15(1), 40–47.
- [56] Ozen, B. F., & Mauer, L. J. 2002. Detection of hazelnut oil adulteration using FT-IR spectroscopy. *Journal of Agricultural Food Chemistry*, 50, 3898–3901.
- [57] Wang, L. Lee, F.S.C. Wang, X. He, Y. 2006. Feasibility study of quantifying and discriminating soybean oil adulteration in camellia oils by attenuated total reflectance MIR and fiber optic diffuse reflectance NIR *Food Chem.* 95, 529.
- [58] Yang, H. Irudayaraj, J. Paradkar, M.M. 2005. Discriminant analysis of edible oils and fats by FTIR, FT-NIR and FT-Raman spectroscopy *Food Chem.* 93, 25.
- [59] Jim'enez Marquez, A. MolinaD'iaz, M.I. 2005. Pascual Reguera, *Sens. Actuator B: Chem.* 107 ,64.
- [60] Sadeghi H.J., Hendra P.J., Wilson R.H., Belton P.S. 1990, Determination of the total unsaturation in oils and margarines by Fourier Transform Raman Spectroscopy. *J. Am. Oil Chem. Soc.* 67 (8), 156-162.
- [61] O'connor R.T. 1956. Application of infrared spectrophotometry to fatty acid derivatives. *J. Am. Oil Chem. Soc.*, 33, 1-15.

[62] Bertrand, D. & Dufour, E. 2006. La spectroscopie infrarouge et ses applications analytiques, 566 (TEC&DOC, Paris,).

[61] Pavia, D. L., Lampman, G. M. & Kriz, G. S. 1996. Introduction to spectroscopy, 416 (Saunders College Publishing,)

Chapitre 4 : LA CHIMIOMETRIE ET SES APPLICATIONS

1. Définition de la Chimiométrie

La chimiométrie est l'utilisation des outils mathématiques combinés aux outils statistiques, pour optimiser les procédures d'obtention et de traitement de données de la chimie analytique, afin d'extraire le maximum d'informations pertinentes. Cette science fait l'objet de nombreuses revues générales [1-5]. Elle a trois objectifs principaux [6]. Le premier est la description des données sous forme synthétique, c'est le cas des statistiques élémentaires ou des méthodes descriptives comme l'analyse en composantes principales (ACP) [7]. Le second objectif est la prédiction soit de valeurs continues comme avec la régression PLS [8], soit de classe comme avec une analyse discriminante [9]. Enfin, son dernier objectif est la planification des expériences [10] pour en tirer le meilleur parti.

La chimiométrie se base sur des modèles qui sont soit linéaires, soit non linéaires. Dans un modèle linéaire, l'hypothèse de départ est que l'information cherchée peut être extraite à partir de combinaisons linéaires des variables de départ, le cas le plus simple étant la régression linéaire [11]. Dans les modèles non-linéaires, le modèle prédictif est sous la forme d'un réseau d'unités connectées entre-elles [6]. L'exemple type de ce modèle est le réseau de neurones [12].

La classification des méthodes chimiométriques peut aussi se faire selon le mode de construction du modèle. On parle de modèle supervisé (ACP, HCA) ou non-supervisées LDA, kNN, SIMCA, Réseaux de neurones à couches.

2. Pré traitements des données spectrales

Les spectres enregistrés peuvent être entachés de défauts que l'on peut tenter de réduire a priori. Les défauts les plus courants sont la présence de bruit aléatoire, la variation de la ligne de base, la redondance de l'information. Ainsi les données spectrales brutes, telles qu'elles sont acquises par un spectromètre, ne revêtent pas obligatoirement la forme la plus adaptée aux traitements chimiométriques ultérieurs. Différents prétraitements peuvent être appliqués aux spectres afin de réduire ces défauts. Parmi les prétraitements les plus utilisés, on trouve la déviation normale standardisée (SNV) (Barnes et al., 1989[13]), la Multiplicative Scatter Correction (MSC), la Correction du Signal Orthogonal OSC (Niazi & Goodarzi, 2008), les lissages et les dérivées (Bertrand & Vigneau, 2006).

- *Standard Normal Variantes (SNV)* [14]

Cette correction améliore la linéarité de la relation entre l'absorbance et la concentration.

Cette méthode est particulièrement intéressante lorsque des techniques de régression linéaire sont utilisées.

Le principal avantage de la SNV est qu'elle s'applique à chaque spectre pris séparément, sans référence, de l'ensemble des échantillons d'étalonnage.

Les données sont centrées et réduites selon l'équation :

$$x_{SNVj} = (x_j - \bar{x}) / \sqrt{\frac{\sum_{k=1}^p (x_k - \bar{x})^2}{p-1}}$$

x_j : valeur de $\log(1/R)$ à la longueur d'onde j
 p : nombre de variables spectrales
 \bar{x} : Valeur moyenne du $\log(1/R)$

Il existe d'autres méthodes de normalisation telles que la MSC (*Multiplicative Scatter Correction*) et la *Weighted MSC*. Pour effectuer la correction MSC, il est nécessaire d'avoir un spectre de référence. Par défaut, le spectre moyen des spectres du lot est utilisé.

De nombreuses études ont montré que les méthodes de normalisation MSC et SNV permettent d'améliorer les résultats de l'analyse quantitative [13,16-17]

- *Les dérivations*

Les dérivations numériques des spectres sont souvent utilisées comme méthodes de prétraitement [14]. La dérivée est historiquement le premier pré-traitement utilisé. Elles permettent de réduire la dérive de la ligne de base, de séparer plus clairement les bandes d'absorption et d'exalter l'information spectrale [18]. Nous pouvons citer comme méthodes, la dérivation par intervalle ou celle basée sur la technique de Savitsky et Golay [19]. Un autre exemple est la dérivée seconde des spectres, largement décrit par D.G. Cameron et al [20].

3. Méthodes descriptives ou exploratoires

3.1 . Analyse en composantes principales (ACP)

L'analyse en composantes principales est la technique d'analyse multivariée la plus utilisée car elle est souvent un point de départ dans un processus d'exploration des données [21,22]. L'analyse en composantes principales (ACP) consiste à rechercher les directions qui représentent le mieux la dispersion des individus dans l'espace multidimensionnel des

variables de départ, et donc à transformer l'espace de représentation des n individus [7]. Par ailleurs, cette technique permet de repérer des éléments présentant des valeurs aberrantes et/ou une structure sous-jacente aux données sans injecter d'information.

Cette technique est utilisée en particulier dans le domaine alimentaire où elle est généralement utilisée conjointement avec l'analyse discriminante.

3.2. Exemples d'applications de l'Analyse en Composantes Principales

Le tableau 7 illustre quelques exemples d'application de l'ACP

Tableau 7 : Principaux travaux analytiques sur divers produits alimentaires, faisant intervenir la ACP ou la LDA ou les deux comme outils de traitement des données (de 1990 à 2002). Liste non exhaustive.

| Aliment(s) | Composés analysés | Technique(s) analytique | Technique(s) chimiométrique(s) | But de l'étude | Année [Réf.] |
|----------------------|--|--|--------------------------------|---|------------------------|
| Fromages | Composés solubles dans l'eau | HPLC | PCA, LDA | Classification | 1990 [23] |
| Huiles | Tous les composés entre 4800 et 800 cm^{-1} | FTIR (Mid-IR) ¹ | PCA, LDA | Authentification | 1994 [24] |
| Purées de fruits | Tous les composés entre 4000 et 800 cm^{-1} | FTIR (Mid-IR) ¹ | PCA, LDA | Authentification | 1995 [25] |
| Jus d'orange | Tous les composés entre 9000 et 4000 cm^{-1} | FTIR (NIR) ¹ | PCA, LDA | Authentification | 1995 [26] |
| Café vert | Tous les composés entre 4800 et 350 cm^{-1} | FTIR (Mid-IR) ¹ | PCA, LDA | Origine | 1996 [27] |
| Huile d'olive vierge | Tous les composés entre 3250 et 100 cm^{-1} | FT-Raman (Far- et Mid-Raman) ¹ | PCR, LDA, HCA | Authentification | 1996 [28] |
| Général ² | Général | FTIR (Mid-IR) | PCA, LDA + autres | Classification et authentification | 1998 [29] |
| Amandes | Acides gras | GC | PCA PCA, LDA | Origine | 1996 [30] 1998 [31] |
| Pdts à base d'ail | Composés soufrés volatils | GC-MS | PCA | Classification | 1998 [32] |
| Miels | Paramètres physico-chimiques | Physico-chimiques et HPLC pour les sucres | LDA | Classification | 1998 [33] |
| Jus de pomme | Arômes | GC-MS capillaire + GC-MS chirale | PCA | Authentification | 1999 [34] |
| Viande | Tous les composés entre 25000 et 4000 cm^{-1} | IR (NIR+Visible) | PCA + autres | Authentification | 2000 [35] |
| Cafés | Paramètres physico-chimiques + Acide chlorogénique | Physico-chimiques et HPLC pour l'acide chlorogénique | PCA | Classification selon l'origine botanique et autres critères | 2001 [36] |
| - | Données de diverses substances de synthèse | MS + IR | PCA, HCA | Comparaison de méthodes de classification | 2001[37] |
| Miels | Sucres | GC-MS | PCA, LDA | Classification selon origine florale | 2001 [38] |
| Vins rouges | Composés phénoliques | HPLC | PCA (+PLS) | Relation composés phénoliques et pouvoir anti-oxydant | 2001 [49] |
| Miels | Tous les composés entre 9000-4000 cm^{-1} | IR (NIR) | PCA, LDA | Classification | 2002 [40] |

4. Les méthodes de classement

L'analyse discriminante est une technique statistique qui permet de discriminer entre des observations compte tenu de leurs caractéristiques individuelles. Son application

empirique a commencé depuis les années 1930 avec les travaux de Fisher [41] et Mahalanobis (1936) [42].

4.1. Classification ascendante hiérarchique

Les méthodes de classification ont pour objectif de regrouper des objets en un nombre restreint de classes homogènes [43]. Pour cela, la classification ascendante hiérarchique (CAH) procède à des regroupements par étapes successives. A chaque étape, les deux classes ou les échantillons, les plus proches, sont combinés pour former une nouvelle classe. Cette méthode produit une séquence complète de partition. Au sommet, on retrouve un grand nombre de classes et, à la base, l'individu étudié est représenté dans une classe unique. Les résultats peuvent alors être mis sous la forme de graphes, appelés dendrogrammes. Les individus sont donc regroupés selon leurs similitudes ou leurs dissimilarités.

Le critère le plus couramment utilisé en classification ascendante hiérarchique est le critère de Ward. A chaque étape, nous cherchons à minimiser l'inertie intra-classe (ou, de manière équivalente, maximiser l'inertie inter-classes). Le critère d'agrégation est, par conséquent, égal à la perte d'inertie inter-classes résultant du regroupement de deux classes. La méthode de Ward a tendance à fournir des classes sphériques et équilibrées en effectifs ce qui n'est pas toujours conforme à la structure des données analysées.

Validation et prédiction

Pour valider un modèle de classification ou prédire l'appartenance d'un échantillon à un groupe, une série d'échantillons est utilisée pour calculer le modèle prédictif. Le barycentre de chaque groupe d'échantillons connus est ensuite calculé et la distance entre chaque nouvel échantillon à classer et les différents barycentres est calculée. Le groupe le plus proche de l'échantillon lui est alors attribué.

Lors de cette validation du modèle, la comparaison des groupes attribués et des groupes réels permet de caractériser le modèle par un taux de bons classements.

4.2 CLV (Clustering of variables around Latent Variables)

Dans un spectre, des variables contiguës ou non, correspondent à des signaux provenant de mêmes molécules ou de mêmes mélanges et varient donc de façon semblable. Pour tenir compte de cette propriété, les variables peuvent être regroupées selon un critère de covariance ou de corrélation.

La technique CLV est très appropriée puisqu'elle a pour but de regrouper des variables autour de variables latentes selon leur covariance ou corrélation carrée. Elle fait une classification hiérarchique des variables en regroupant les variables maximisant le paramètre

T qui prend en compte l'un ou l'autre des critères. Les variables latentes sont définies comme la première composante principale de chaque groupe établi.

Afin de déterminer le nombre de groupes le plus représentatif des données, le suivi de l'évolution du critère T est déterminant. En effet, si la variation de T est importante lorsque le nombre de groupes diminue cela signifie qu'une part importante de l'information contenue dans la partition précédente a été perdue.

5. Méthodes prédictives

5.1 Régression linéaire et Méthode des moindres carrés

C'est une des méthodes les plus couramment appliquée en spectroscopie. La régression PLS réduit la matrice. Cependant, le choix des nouveaux axes ne se fait plus en fonction de la variance expliquée maximale, comme dans l'APC, mais selon les directions les plus pertinentes en termes de prédiction des variables dépendantes, c'est-à-dire en fonction des covariances maximales entre les variables indépendantes et dépendantes.

L'estimation des coefficients est réalisée par régression linéaire au sens des moindres carrés. Cette méthode initiale consiste à mesurer l'efficacité d'un ajustement par la somme des carrés des écarts r_i (écarts appelés "résidus") entre la valeur vraie de la concentration en analyte et sa valeur estimée par la courbe d'étalonnage. Faire un ajustement c'est minimiser la "distance "

Dans le cas de l'analyse d'un seul composé à partir d'une seule bande d'absorption, la concentration n'est liée qu'à une seule variable, l'absorbance de la bande considérée :
[Analyte] = a Absorbance + b.

La régression linéaire multiple et l'analyse multicomposants

Toute méthode d'analyse quantitative multicomposant classique repose donc sur la régression linéaire multiple or deux obstacles s'opposent à l'utilisation de cette technique dans le cas de l'analyse spectrale : le grand nombre de variables explicatives X_i et leurs corrélations.

Le nombre $X_i Y_i$: lorsqu'on analyse un mélange de composés, aucune absorbance à une longueur d'onde déterminée n'apporte une information suffisante en soi sur l'un des constituants du mélange. En effet, dans le cas général, plusieurs composés si ce n'est tous, absorbent plus ou moins à une fréquence donnée. Pour obtenir une information suffisante, précise et pertinente, il est absolument nécessaire d'utiliser les absorbances d'un grand nombre de longueurs d'ondes (le maximum théorique étant chaque point d'échantillonnage du spectre), ce sont ces absorbances qui sont les variables explicatives X_i de la concentration en analyte Y_i recherchée.

Les corrélations : pour un analyte et une bande d'absorption donnés, les variations d'un grand nombre d'absorbances successives sont naturellement **fortement corrélées** entre elles puisque quand l'intensité de cette bande croît ou décroît, les absorbances de deux points voisins varient de la même manière (et ceci, à intervalles répétés dans le spectre !).

Compression des données : pour surmonter ces deux contraintes, nous devons utiliser une technique mathématique qui nous permette de condenser **toute** l'information expérimentale en un nombre plus réduit de **nouvelles** variables explicatives **non corrélées** entre-elles.

Les deux versions de la PLS, PLS1 et PLS2, présentent de petites différences, mais très importantes. Pour les cas où l'on veut modéliser une matrice \mathbf{X} et sa relation avec un vecteur \mathbf{y} , on utilise la PLS1. Si l'on a une matrice \mathbf{Y} et on veut modéliser chacun des vecteurs, on peut prendre une colonne de \mathbf{Y} à la fois et appliquer le PLS1, ce qui donne des résultats plus performants que si on utilise la PLS2. Celle-ci est utilisée quand on veut modéliser toute une matrice \mathbf{Y} d'un seul coup, mais la décomposition n'est pas optimisée pour un \mathbf{y} particulier, ce qui peut diminuer la qualité des prédictions pour de nouveaux échantillons.

5.2 iPLS

iPLS [44,45] est une méthode de sélection de variables appliquant l'algorithme PLS sur de petites zones de spectre. Il s'agit soit de définir le nombre d'intervalles souhaités, ce qui déterminera la taille des zones spectrales et donc le nombre de variables à prendre en compte, soit de définir manuellement les intervalles à étudier. Sur chacun de ces intervalles, un modèle PLS est calculé avec un nombre croissant de variables latentes jusqu'à un nombre maximal, f , défini par l'utilisateur. Il faut aussi introduire un vecteur \mathbf{y} de valeurs à prédire par régression PLS. Le modèle de référence utilise la totalité du spectre pour la prédiction avec une dimensionnalité déterminée par validation-croisée.

Pour chaque zone spectrale on obtient f valeurs de RMSECV qu'il faut alors comparer à la valeur obtenue sur les spectres entiers, RMSECV_{réf} avec un nombre de variables latentes, déterminé par cross-validation .

Méthodes descriptives multidimensionnelles

Lorsqu'un ensemble de m variables aléatoires doit être réduit pour en simplifier l'analyse, les p premiers axes de l'analyse en composantes principales ou indépendantes, correspond à un meilleur choix de représentation.

5.3 .Exemples d'applications de la méthode des moindres carrés partiels en analyse quantitative

Les exemples rapportés par M.Tenenhaus [46] illustrent les avantages de la régression PLS par rapport aux méthodes anciennes, la variété des domaines d'application et l'efficacité de cette méthode de régression dans le cas des données spectroscopiques. De nombreuses études d'analyse quantitative récemment publiées dans différents domaines (détergence, agroalimentaire, médecine...) utilisent la méthode des moindres carrés partiels [47-48-54]. Dans le cas précis de l'analyse quantitative infrarouge nous avons retenu les travaux suivants [55-60].

C'est le domaine des études socio-économiques qui connaît les premiers développements de la méthode des moindres carrés partiels (PLS). L'algorithme NIPALS proposé par H.Wold [61] est à l'origine du développement de la régression PLS. Son utilisation remonte aux années 80 avec les travaux de S.Wold et coll. [62-63]. La méthode PLS est depuis lentement reconnue dans le domaine des industries chimiques, pétrolières et agroalimentaires [64].

Un grand nombre de travaux et d'ouvrages décrivent la méthode PLS. On retiendra ceux de S.Wold et coll. [65], de K.Esbensen et coll. (« Multivariate Analysis in Practice » (1994)), de A.Hoskuldsson (« Prediction Methods in science and Technology, Vol 1 Basic Theory » 1996). Pour aborder l'utilisation de la méthode PLS en spectroscopie infrarouge quantitative, nous retenons particulièrement l'ouvrage de H.Martens et T.Naes consacré à la Calibration multivariée [66] et la mise au point de M.P.Fuller et coll. [47].

Comme nous l'avons exposé dans le cadre général de la régression linéaire, cette méthode prédictive est construite spécifiquement pour l'étude de la compréhension et de la description des relations souvent très complexe entre X et Y en l'absence d'un modèle théorique.

Elle prend en compte la multicollinéarité, ce que ne fait pas la régression multiple classique. Cette méthode reprend la philosophie de base de la régression en composantes principales mais, dans la méthode PLS, la décomposition en composantes principales est réalisée **simultanément** et de manière **interdépendante** entre les **deux** tableaux X et Y : la matrice Y sert de "guide" à la décomposition de X. Reprenons sous forme schématique les différentes matrices que l'on manipule :

R₂ : Les données sont centrées et réduites selon l'équation

$$R^2 = \frac{\sum_{i=1}^n (\hat{Y}_i - Y)^2}{\sum_{i=1}^n (Y - \bar{Y}_i)^2}$$

Y_i : La concentration observée

\hat{Y}_i : La concentration prédite par le modèle de régression

\bar{Y} : La concentration moyenne observée.

$$\text{RMSEC \%} = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_{ic})^2}{n}} \times \frac{100}{\bar{Y}}$$

\hat{Y}_{ic} : La concentration prédite par PLS

pour un échantillon d'étalonnage

$$\text{RMSECV \%} = \sqrt{\frac{(Y_i - \hat{Y}_{icv})^2}{n}} \times \frac{100}{\bar{Y}}$$

Validation croisée ou cross validation

La validation croisée ou cross validation est une méthode de validation interne de modèles.

C'est un critère permettant d'évaluer la consistance interne des données.

Elle consiste à enlever un étalon de la matrice d'étalonnage de taille m et à élaborer le modèle PLS avec les $m-1$ étalons restants. Puis, on applique le modèle à l'étalon rejeté. On répète l'opération pour les m étalons de la matrice. Chaque étalon possède alors une valeur réelle et une valeur estimée. L'écart entre ces deux valeurs est appelé erreur ou résidu et noté r_i . Le calcul de la somme des carrés des résidus permet de définir le PRESS.

L'offre logicielle

Le marché des logiciels scientifiques produit des logiciels de plus en plus conviviaux et puissants. Parmi les logiciels généralistes de statistiques utilisant l'approche de la méthode PLS, on peut citer : SIMCA, UNSCRAMBLER, LUMIERE, STATBOX.

Dans notre étude nous avons travaillé avec les versions d'UNSCRAMBLER 9.8 et 10.1

Il existe des manuels d'utilisation des outils chimiométriques pour l'analyse quantitative de spectres [67, 68].

5.4 PLS DA

L'analyse discriminante PLS (PLSDA) est une méthode de classification supervisée combinant la régression PLS et l'analyse discriminante linéaire. Elle est très couramment utilisée dans le domaine de l'agroalimentaire [70].

Dans le cas de PLS-DA, les variables Y ne sont plus quantitatives, mais qualitatives (variables binaires 0/1).

La PLS2 trouve aussi une application énorme dans l'analyse discriminante, où elle est plus connue sous le nom de PLS-DA. En fait, dans la PLS-DA on applique la PLS2 à une

matrice \mathbf{Y} contenant des colonnes de valeurs binaires pour le classement des échantillons dans des groupes (1 si l'objet appartient au groupe représenté par cette colonne et 0 dans le cas contraire).

Il existe de nombreuses variations et extensions de la méthode PLS standard, par exemples la «PLS-Discriminant Analysis» (PLS-DA) [70-72] pour l'analyse discriminante basée sur une matrice \mathbf{Y} de valeurs binaires d'appartenance, «PLS-cluster» [73] pour la classification hiérarchique prédictive,

L'analyse des clusters dans un jeu de données est très utilisée dans le domaine de la chimométrie. Cette technique d'exploration des données permet d'étudier les groupements dans les données, et les relations entre les objets (les échantillons).

PLS-Cluster [73-74] est une méthode récursive qui a pour but de regrouper les individus ayant des caractéristiques semblables, sans *a priori* sur les échantillons et leur véritable proximité.

La première étape de PLS-Cluster sépare les échantillons en G groupes, selon la méthode décrite au paragraphe suivant. Chacun de ces groupes est lui-même soumis à l'algorithme PLS-Cluster, pour être séparé en G_i groupes. La procédure se poursuit jusqu'à ce que chaque échantillon soit dans un singleton et que l'on obtienne ainsi une hiérarchie. A chaque étape de cette classification hiérarchique, le nombre de subdivisions à chaque noeud est déterminé à partir de la structure interne des données.

A partir de la matrice de données (par exemple une matrice spectrale \mathbf{X}), on crée une variable dépendante pouvant prendre une valeur comprise entre 0 et 1 (un vecteur \mathbf{y}), de façon à pouvoir construire le modèle PLS qui permettra de grouper les échantillons. Dans l'algorithme initial, cette variable dépendante était générée de façon aléatoire [73] et cette variable ne pouvait prendre que les valeurs 0 et 1. Dans l'algorithme généralisé, cette variable est initialisée comme étant le premier vecteur "score" normé de la matrice \mathbf{X} , ce qui permet de mieux tenir compte de la structure des données [74]. Un modèle PLS est construit pour relier \mathbf{y} à \mathbf{X} , et grâce à ce modèle, on peut déterminer des valeurs prédites ("ajustées") \mathbf{y}_p . On teste la convergence entre \mathbf{y} et \mathbf{y}_p : tant qu'il n'y a pas convergence, on remplace \mathbf{y} par \mathbf{y}_p , et on recalcule un modèle PLS. Une fois la convergence atteinte, les valeurs sont ordonnées et les différences entre valeurs consécutives sont calculées pour déterminer le nombre de groupes présents.

La validation croisée est une méthode de validation interne de modèles. C'est un critère permettant d'évaluer la consistance interne des données.

6. Validation-croisée

a. Objectif: *détermination du nombre optimal de variables latentes*

D'une certaine façon, la régression PLS comme d'autres méthodes de projection des données, donne des approximations d'une façon similaire aux polynômes de Taylor. On peut construire des modèles de plus en plus proches des données simplement en augmentant le nombre de variables latentes utilisées. Ceci entraîne une diminution des résidus avec l'augmentation du nombre de variables latentes (ou des composantes principales dans le cas de la PCR). Dans le cas de l'utilisation de la PLS comme méthode de régression, on n'a pas comme objectif la diminution des écarts par rapport à l'étalonnage mais par rapport à la prédiction. On veut s'approcher le plus des valeurs de la variable réponse, mais on veut utiliser seulement l'information pertinente dans le jeu de données d'étalonnage pour construire un modèle de prédiction qui minimise les écarts pour de nouveaux échantillons. Tout ce qui est variabilité intrinsèque (bruit) des échantillons et qui n'a pas de rapport avec la variance intéressante pour la description des phénomènes étudiés ne doit pas être inclus dans le modèle, sous peine de sur-ajustement («overfitting») du modèle. Cet effet correspond à une augmentation de l'exactitude et de la précision de l'étalonnage, mais une diminution pour les prédictions. Le sous-ajustement correspond à un modèle avec un nombre de variables latentes inférieure à ce qu'il faut pour inclure toute la variabilité nécessaire pour minimiser les erreurs de prédiction.

La méthode PLS est donc fortement dépendent d'un judicieux choix du nombre de variables latentes à inclure dans le modèle, ce qui implique un équilibre entre l'extraction de la variabilité des matrices (ajustement) et un bon pouvoir prédictif. Elle a été initialement proposée par Stone (1974) pour l'ACP (Analyse en composantes multiples).

Plusieurs méthodes existent pour choisir le nombre de variables d'un modèle de régression et pour évaluer l'incertitude des modèles. La plus utilisée, la validation croisée («cross-validation»), estime l'exactitude des prédictions pour des modèles avec un nombre de variables latentes croissant.

Il y a plusieurs variantes de la validation croisée par rapport à différents facteurs, comme le nombre d'objets faisant partie des groupes d'étalonnage/validation. Pour simplifier, on considère le cas d'une seule variable réponse, donc $q = 1$ et chaque groupe de validation ne contient qu'un seul échantillon (procédure «leave-one-out»), donc $m = 1$.

Références bibliographiques

- [1] Brereton, R. G. 1990, Chemometrics applications of mathematics and statistics to laboratory systems, 307 (Ellis Horwood, Chichester).
- [2] Beebe, K., Pell, R. & Seasholtz, M. 1998, Chemometrics: A practical Guide, 360 (Wiley, New York).
- [3] Eriksson, L., Johansson, E., Kettaneh-Wold, N. & Wold, S. 2001, Multi- and Megavariate Data Analysis - Principles and Applications, 533 (UMETRICS).
- [4] Martens, H. & Naes, T. Multivariate calibration, 438 (Wiley, New York, 1989).
- [5] Sharaf, M. A., Illman, D. L. & Kowalski, B. R. 1986, Chemometrics, 352 (Wiley, New York,).
- [6] Bertrand, D. & Dufour, E. 2006, La spectroscopie infrarouge et ses applications analytiques, 566 (TEC&DOC, Paris).
- [7] Joliffe, I. T. 1986, Principal Component Analysis, 271 (Springer-Verlag, New York.
- [8] Tenenhaus, M. 1999, L'approche PLS. Revue de Statistique Appliquée 47, 2-55.
- [9] Bardos, M. 2001 Analyse Discriminante (Application au risque et scoring financier), 223 (Dunod, Paris).
- [10] Goupy, J. 1996, La Méthode des Plans d'Expériences, 303 (Dunod, Paris,).
- [11] Daudin, J.-J., Robin, S. & Vuillet, C. 1999, Statistique inférentielle, 185 (Presses Universitaires de Rennes, Rennes).
- [12] Haykin, S. 1998, Neural Networks: A Comprehensive Foundation, 842 (Prentice Hall, New Jersey).
- [13] Barnes, R. J., Dhanoa, M. S., Lister, S. J., 1989, Standard normal variate transformation and detrending of near-infrared diffuse reflectance spectra. Applied Spectroscopy 45, 772-777.
- [14] Bertrand D. E. M. Qannari, 2000, in : La spectroscopie infrarouge et ses applications analytiques, TEC&DOC, Londres-Paris-New York, 251–279.
- [15] Laloum E. 2004, Une méthode chimiométrique originale d'identification de produits par spectroscopie proche infrarouge Spectra. Analyse. , 33, 34–36.
- [16] Manley, M. McGill A et Osborne B., 1994, J .Near Infrared Spectrosc., 2, 93.

- [17] Dhanoa, M.S., S.J. Lister, R. Sanderson, and R.J. Barnes, 1994 .The Link Between Multiplicative Scatter Correction (MSC) and Standard Normalariate (SNV) Transformations of NIR Spectra, *J. Near Infrared Spectrosc.* 2,43–47.
- [18] Roggo, Y. 2003 : Détermination de la qualité de la betterave sucrière par spectroscopie proche Infrarouge et chimiométrie. Thèse de doctorat, Univerisité de Sciences et Technologies de Lille.
- [19] Savitsky A.et Golay M., 1964, Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.*, 36 (8), 1627–1639
- [20] Cameron D.G., Moffatt D.J., *Applied Spectroscopy*, (1987) Vol. 41, N°4.
- [21] Leardi, R. 2002, *Chemometrics: From Classical to Genetic Algorithms*. *Grasas y Aceites*, 53(1), 115-127.
- [22] Siebert, K. J. 2001, *Chemometrics in Brewing – A Review*. *J. Am. Soc. Brew. Chem.*, 59(4), 147-156.
- [23] Molher-Smith, A. Nakai, S. 1990, Classification of Cheese Varieties by Multivariate Analysis of HPLC Profiles. *Can. Inst. Food Sci. Technol. J.*, 23(1), 53-58.
- [24] Lai, Y. W.; Kemsley, E. K.; Wilson, R. H. 1994, Potential of Fourier Transform Infrared Spectroscopy for the Authentication of Vegetable Oils. *J. Agric. Food Chem.* , 42, 1154-1159.
- [25] Defernez, M.; Kemsley, E. K.; Wilson, R. H. 1995, Use of Infrared Spectroscopy and Chemometrics for the Authentification of Fruit Purees. *J. Agric. Food Chem.*, 43, 109-113.
- [26] Twomey, M.; Downey, G.; McNulty, P. B. 1995. The Potential of NIR Spectroscopy for the Detection of the Adulteration of Orange Juice. *J. Sci .Food Agric.*, 67, 77-84.
- [27] Suchanek, M.; Filipova, H.; Volka, K.; Delgadillo, I.; Davies, A. N. 1996, Qualitative Analysis of Green Coffee by Infrared Spectroscopy. *Fresenius J. Anal. Chem.*, 354, 327-332.
- [28] Baeten, V.; Meurens, M. 1996, Detection of Virgin Olive Oil Adulteration by Fourier Transform Raman Spectroscopy. *J. Agric. Food Chem.*, 44, 2225-2230.
- [29] Downey, G. 1998, Food and Food Ingredient Authentication by Mid-Infrared Spectroscopy and Chemometrics. *Trends in Anal. Chem.* , 17(7), 418-425.
- [30] García-López, C.; Grané-Teruel, N.; Berenguer-Navarro, V.; García-García, J. E.; Martín-Carratalá, M. L. Major 1996, Fatty Acid Composition of 19 Almond Cultivars of Different Origins. A Chemometric Approach. *J. Agric. Food Chem.*, 44, 1751-1755.
- [31] Martín-Carratalá, M. L.; García-López, C.; Berenguer-Navarro, V.; Grané-Teruel, N. 1998, New Contribution to the Chemometric Characterization of Almond Cultivars on the Basis of Their Fatty Acid Profiles. *J. Agric. Food Chem.*, 46, 963-967.

- [32] Mochizuki, E.; Yamamoto, T. 1998, Identification of Allium Products Using Flame Photometric Detection Gas Chromatography and Distribution Patterns of Volatile Sulfur Compounds. *J. Agric. Food Chem.*, 46, 5170-5176.
- [33] Mangas, J. J.; Moreno, J.; Picinelli, A.; Blanco, D. 1998, Characterization of Cider Apple Fruits According to Their Degree of Ripening. A Chemometric Approach. *J. Agric. Food Chem.*, 46, 4174-4178.
- [34] König, T.; Schreier, P. 1999, Application of Multivariate Statistical Methods to Extend the Authenticity Control of Flavour Constituents of Apple Juice. *Z. Lebensm Unters Forsch A*, 208, 130-133.
- [35] Ding, H. B.; Xu, R. J. 2000, Near-Infrared Spectroscopy Technique for Detection of Beef Hamburger Adulteration. *J. Agric. Food Chem.*, 48, 2193-2198.
- [36] Maeztu, L.; Andueza, S.; Ibañez, C.; Paz de Peña, M.; Bello, J.; Cid, C. 2001, Multivariate Methods for Characterization and Classification of Espresso Coffees from Different Botanical Varieties and Types Roast by Foam, Taste, and Mouthfeel. *J. Agric. Food Chem.*, 49, 4743-4747.
- [37] Schoonjans, V.; Massart, D.L. 2001, Combining Spectroscopic Data (MS, IR): Exploratory Chemometric Analysis for Characterizing similarity/diversity of Chemical Structures. *J. Pharm. Biomed. Anal.*, 26, 225-239.
- [38] Terrab, A.; Vega-Pérez, J.M.; Díez, M.J.; Heredia, F.J. 2001, Characterization of Northwest Moroccan Honeys by Gas Chromatographic-Mass Spectrometry Analysis of Their Sugar Components. *J. Sci. Food Agric.*, 82, 179-185.
- [39] Brenna, O.V.; Pagliarini, E. 2001, Multivariate Analysis of Antioxidant Power and Polyphenolic Composition in Red Wines. *J. Agric. Food Chem.*, 49, 4841-4844.
- [40] Davies, A.M.C.; Radovic, B.; Fearn, T.; Anklam, E. A 2002, Preliminary Study on the Characterization of Honey by Near Infrared Spectroscopy. *J. Near Infrared Spectrosc.*, 10, 121-135.
- [41] Fisher, R.A. 1936, "The use of multiple measurements in taxonomic problems", *Annals of Eugenics*, 7, 179-18.
- [42] Mahalanobis, P.C. 1963, "On the generalized distance in statistics", *Natl. Inst. Science*, 12, 49-55.
- [43] Lebart L., Morineau A., Piron M. 2006: *Statistiques exploratoires multidimensionnelles*, 4ème édition, Dunod .

- [44] Norgaard, L. et al. 2000, Interval Partial Least-Squares Regression (iPLS): A Comparative Chemometric Study with an Example from Near-Infrared Spectroscopy. *Applied Spectroscopy* 54, 88A-113A .
- [45] Borin, A. & Poppi, R. J. 2005, Application of mid infrared spectroscopy and iPLS for the quantification of contaminants in lubricating oil. *Vibrational Spectroscopy* ,37, 27-32.
- [46] Tenenhaus M. La regression PLS, théorie et pratique Paris : edition TECHNIP, 1998, 254 .
- [47] Fuller M.P., Ritter G.L., and Drapper C.S. 1988, Partial Least-squares Quantitative Analysis of Infrared Spectroscopic Data: Part I: Algorithm Implementation et part II: Application to detergent Analysis *Applied Spectroscopy*, 42, n° 2, 217-236
- [48] Lee K.R., Zuber G., Katrincic L. 1988, Chemometrics approach to the determination of polymorphism of a drug compound by infrared spectroscopy *Drug development and industrial pharmacy*, X, 26, n°2, 135-147
- [49] Liu K.Z., Mantsch H.H. 1999, Simultaneous quantification from infrared spectra of glucose concentrations, lactate concentrations, and lecithin/sphingomyelin ratios in amniotic fluid *American journal of obstetrics and gynecology*, 180, n°3, part 1, 696-702
- [50] Carpignano R., DE Giorgi M.R. 1999, Structure optimization in a series of azo dyes by experimental design and PLS modeling *Quimica Analitica*, 18, n°1, 99-111
- [51] Revilla A.L., Vargas G., Havel J. 1999, Spectrophotometric analysis of alpha- and beta-naphthol mixtures applying partial leastsquares method *Chemical Papers Chemické Zvesti*, 53, n°1, 34-39
- [52] Duponchel Ludovic, Huvenne Jean Pierre (dir) 1997, Validation par les méthodes chimiométriques d'un analyseur spectrométrique de lait travaillant dans le proche infrarouge. Transfert de calibration en vue de développements industriels thèse de l'université de Lille 1/ Villeneuve – d'Ascq/FRA, 210p
- [53] Bouhsain Z., Garrigues S., De La Guardia M. 1996, Simultaneous stopped-flow determination of paracetamol, acetylsalicylic acid and caffeine in pharmaceutical formulations by Fourier transform infrared spectrometrywith partial leastsquares data treatment *Analyst*, 121, n°12, 1935-1938
- [54] International conference of fourier transform spectroscopy (8th) 1991, Lubeck Travemunde, Allemagne) HAALAND D.M. Multivariate calibration applied to the quantitative analysis of infrared spectra *Allemagne*, 19.

- [55] Dupuy N., Wojciechowski C., Huvenne J.P. 1999, Quantitative analysis of edible trans fats by Fourier Transform Infrared spectroscopy from raw materials to finished products *Sciences des aliments*, 19, n°6, 677-686
- [56] Sundqvist S., Leppamaki M., Paatero E., Minkkinen P. 1999, Application of IR spectroscopy and multivariate calibration to monitor the fusion synthesis of Ca- and Ca/Mg-resinates *Analytica Chimica Acta*, , 391, n°3, 269-276
- [57] Man Y.B.C., Mo M.H., Van De Voort F.R. 1999, Determination of free fatty acids in crude palm oil and refined-bleached-deodorized palm olein using Fourier transform infrared spectroscopy *Journal of the American Oil Chemists Society*, X, 76, n°4, 485-490
- [58] Schindler R., Lendl B. 1999, Simultaneous determination of enzyme activities by FTIR spectroscopy in an on-step assay *Analytica chimica acta*, 391, n°1, 19-28
- [59] Yan Li, Junde Wang, Weiqun Yuan 2000, Simultaneous determination and spectra correction of air toxic organic compounds using PLS and ANN in FTIR spectroscopy *Journal of environmental science and health, part A : Toxic/Hazardous Substances & Environmental Engineering*, , 35, n°9, 1673-1691
- [60] Mattu M.J., Small G.W., Combs R.J., Knapp R.B., Kroutil R.T. 2000, Quantitative analysis of sulfur dioxide with passive Fourier transform infrared remote sensing interferogram data *Applied spectroscopy*, 54, n°3, 341-348
- [61] Wold H. *Multivariate Analysis* New York, Academic press, 1966, p 391-420
- [62] Lindberg W., Person J., Wold S. 1983, Partial Least Square Method for Spectrofluorimetric Analysis of Mixtures of Humic acid and Ligninsulfonate *Anal Chem*, 55, 643-648
- [63] Sjoström M., Wold S., Lindberg W., Person J., Martens H., 1983, *Anal chim Acta*, 150, 61
- [64] Martens H., Jensen S.A. 1983, *Progress in Cereal Chemistry and technology, Proceeding, 7th World cereal and bread Congress*, vol 5a HOLAS J. et KRATOCHVIL J., Eds (Elsevier Amsterdam)
- [65] Wold S., Martens H., Wold H. 1982, The multivariate calibration problem in chemistry solved by PLS method *Proceedings, conference on Matrix Pencils*, RUHE A. et KAGSTROM B., Eds (Heidelberg, 1982), 286-293
- [66] Martens H., NAES T. *Multivariate calibration* New York, John Wiley & Sons, 1989, 419.
- [67] Difoggio Rocco 2000, Guidelines for applying chemometrics to spectra : Feasibility and error propagation *Applied spectroscopy*, 54, n°3, 94A-113A

- [68] Lavine Barry K. 1999, Book reviews : Chemometric techniques for quantitative analysis Applied spectroscopy, 53, n°1
- [69] Downey, G. 2006 in La spectroscopie infrarouge et ses applications analytiques (ed. Bertrand, D.), 479-504 (TEC & DOC, Paris).
- [70] Indahl, U.G.; Martens, H.; Naes, T. 2007. «From dummy regression to prior probabilities in PLS-DA», Journal of Chemometrics 21, 529-536.
- [71] Chevallier, S.; Bertrand, D.; Kohler, A.; Courcoux, P. 2006. «Application of PLS-DA in multivariate image analysis», Journal of Chemometrics, 20, 221-229.
- [72] Barker, M.; Rayens, W.; 2003. «Partial least squares for discrimination», Journal of Chemometrics, 17, 166-173.
- [73] Barros, A.S.; Rutledge, D.N.; 2004, «PLS_Cluster: a novel technique for cluster analysis», Chemometrics and intelligent laboratory systems, 70 (2), 99-112.
- [74] Jouan-Rimbaud Bouveresse, D., Barros, A. S. & Rutledge, D. N. 2007, Generalised PLS_Cluster: an extension of PLS_Cluster for interpretable hierarchical clustering of multivariate data. Sensing and Instrumentation for Food Quality and Safety, 1, 79-90.



Partie Expérimentale

Chapitre 5: CLASSIFICATION DES AMANDES D'ARGAN SELON L'ORIGINE GEOGRAPHIQUE PAR LA SPECTROSCOPIE MOYEN INFRAROUGE COUPLEE A LA CHIMIOMETRIE

1. Introduction

L'arganier *d'argania spinosa*, un arbre endémique du sud-ouest du Maroc [1], bien connu pour ses vertus cosmétiques, pharmaceutiques et alimentaires, Il est considéré comme un produit de luxe dû à ses propriétés uniques [2]. En raison de la valeur élevée du produit, la possibilité de pratiques illégales telles que la dilution avec des huiles meilleures marchées (par exemple, addition d'huile d'olive colorée avec le paprika ou d'autres substances) devrait être contrôlée. Ainsi, il est important d'empêcher la fraude en garantissant son origine et sa qualité, commandant des adultérations possibles.

L'authentification de l'arganier a pris différents aspects, y compris l'adultération, la caractérisation, mauvais étiquetage et l'origine trompeuse. Les Organismes de l'industrie alimentaire et de la réglementation sont obligés de faire un effort de plus en plus d'authentifier l'origine des denrées alimentaires et de contrôler les constituants alimentaires pour assurer la sécurité sanitaire du produit et de répondre aux exigences de qualité des consommateurs [3-5] .

Le développement de nouvelles techniques de plus en plus sophistiquées pour l'authentification des produits alimentaires augmente avec l'augmentation de la conscience du consommateur des issues de sûreté et d'authenticité de nourriture [6]

Beaucoup d'études ont été réalisées afin de trouver le lien entre la composition et l'origine géographique en utilisant les techniques analytiques, en général couplées à l'analyse multivarié. Les techniques chromatographique [7] (chromatographie en phase gazeuse; chromatographie en phase gazeuse couplée à la spectrométrie de masse; et la chromatographie liquide, la chromatographie liquide sous haute pression) ont été intensivement employées pour quantifier les acides gras [8-10] , aussi bien les composés mineurs ont été classés que les stérols [11], les composés phénoliques [12-13], et les hydrocarbures [14-15].

Cependant, l'analyse chromatographique de routine est souvent longue et coûteuse et ne donne pas d'information directe sur la qualité globale de l'échantillon [16]. Par conséquent, les mesures alternatives moins chères, plus rapides et non destructives qui peuvent fournir une réponse rapide sur l'origine de l'échantillon doivent être explorées.

Dans ce sens la spectroscopie moyen infrarouge (MIR) a beaucoup d'applications parce que les empreintes digitales des groupes fonctionnels peuvent être montrées étroitement et intensément dans cette région de MIR (4000-400 cm^{-1}). La spectroscopie infrarouge à transformée de Fourier (FTIR) avec l'accessoire réflexion totale atténuée (ATR) ou de la transmission a été employée pour authentifier, identifier et classifier des graisses et les huiles [17-19].

Le but de notre travail est de développer une nouvelle application basée sur la technique FTIR combinée avec les méthodes supervisées et des méthodes non-supervisées pour déterminer l'origine géographique des amandes d'argan. La vérification de l'origine, ou la détermination de l'origine d'un échantillon non identifié des amandes d'argan est un problème approprié en raison de l'utilisation et de l'exportation croissante de ce produit naturel. Le travail actuel vise à employer le plein profil IR des amandes d'argan en combinaison avec des méthodes multivariées pour la classification des échantillons des amandes d'argan provenant de trois régions marocaines.

2. Matériel et méthodes

2.1. Echantillons

30 échantillons des amandes d'argan ont été collectés à partir de trois régions marocaines au cours de la saison de la récolte 2010 (10 échantillons par région). La méthode de collecte (dépulpage, le concassage des noix et l'obtention des amandes d'argan), les conditions de stockage et de conservation sont les mêmes pour tous les échantillons utilisés. Des échantillons de noyaux d'amandes naturelles ont été analysés, sans être soumis à des traitements de torréfaction. L'enquête a porté sur des échantillons d'argan cueillis dans les zones désignées Ait Baha (AITB), Agadir (AGT) et Guelmim (GLM) (Figure 1). Les caractéristiques géographiques comme l'altitude, la latitude et la longitude ainsi que les précipitations moyennes de ces écorégions sont résumées dans le tableau 8.

Deux lots ont été préparés, un lot de 21 échantillons des amandes d'argan pour la calibration et 9 échantillons des amandes d'argan ont été utilisés pour le lot de validation

Tableau 8 : Les données climatiques des trois régions étudiées.

| | Latitude N | Longitude w | Altitude (m) | précipitation (mm) |
|----------|------------|-------------|--------------|--------------------|
| Agadir | 30°41' | 9°33' | 150–350 | 300 |
| Guelmim | 28°58' | 10°3' | 286 | 217 |
| Ait baha | 29°33' | 8°58' | 340-1380 | 218 |

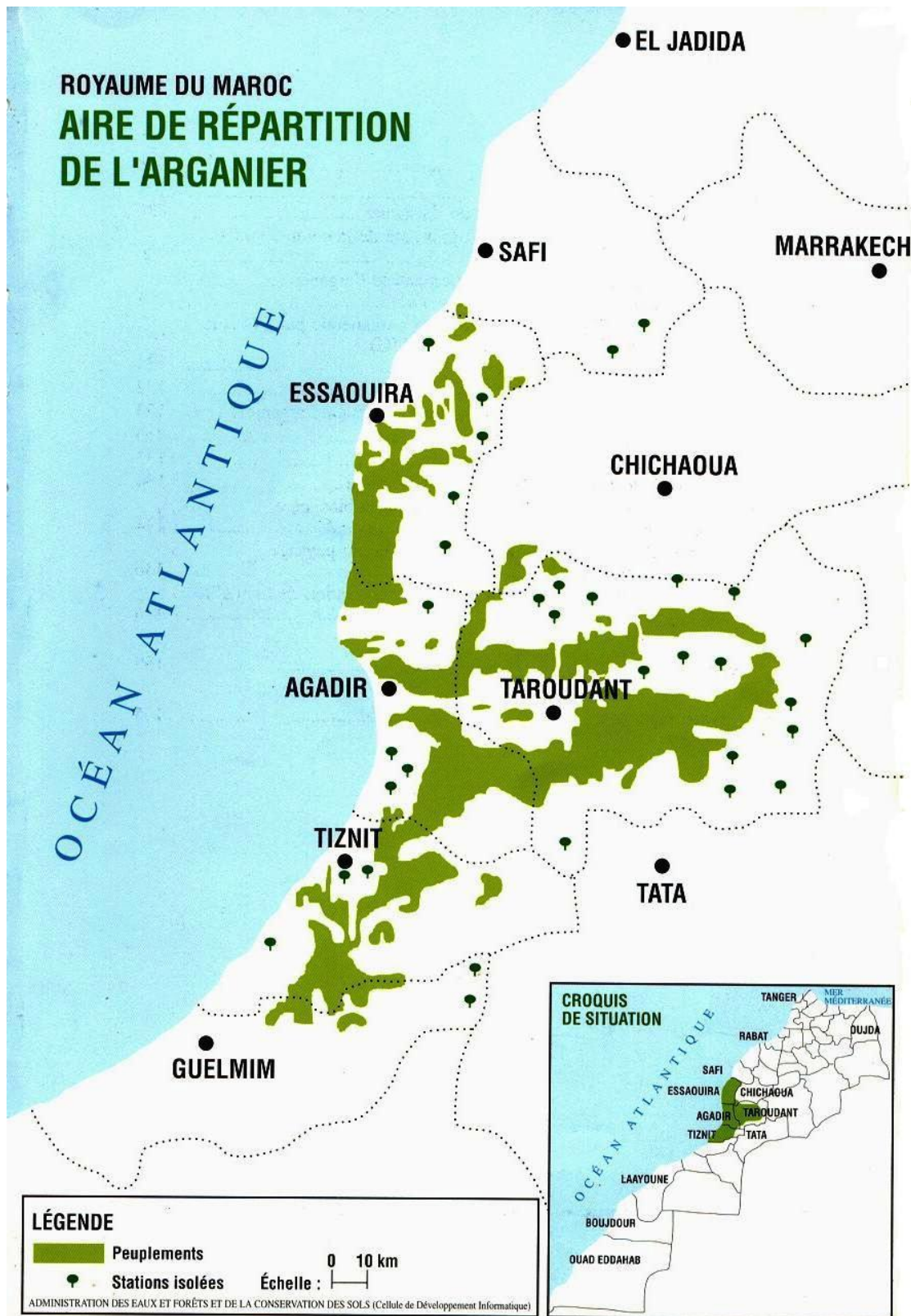


Figure 1 : l'aire de répartition de l'arganier dans le sud-ouest marocain [20]

2.2. Spectroscopie Moyen Infrarouge (MIR)

Le spectromètre IR à Transformée de Fourier utilisé pour cette étude est le Bruker (Vector 22) permettant l'acquisition de spectres dans le moyen infrarouge en fonction de la combinaison source/séparatrice/détecteur utilisée.

Cet instrument est équipé d'un accessoire total atténué (ATR) (cristal diamant) et d'un détecteur contenant du deutérium du sulfate de triglycine (DTGS). Le cristal d'ATR a été soigneusement nettoyé avec de l'éthanol pur pour éliminer la présence des résidus des amandes d'argan entre les mesures pour assurer une surface en cristal propre pour obtenir les meilleurs spectres d'échantillon. Les échantillons ont été analysés juste après la collecte.

Le Vector 22 est piloté par un logiciel d'acquisition et de traitement de données "OPUS version 5.0.

Le spectre de chaque échantillon a été enregistré dans la gamme spectrale de 4000 à 600 cm^{-1} . Une somme totale d'un nombre de scan de 104 a été rassemblée et ramenée à une moyenne par chaque échantillon. La résolution spectrale était le cm^{-1} . La version 5.5 de logiciel d'OPUS a été utilisée pour la collecte de données spectrale.

2.3. Analyse statistique des données

Les différents prétraitements testés sur les données spectrales FTIR collectées sont: la déviation normale standardisée (SNV), la première et la seconde dérivée avec un nombre de polynômes différent.

L'analyse hiérarchique ascendante est premièrement appliquée sur le lot de calibration pour détecter la possibilité de distinguer entre les trois groupes de régions des amandes d'argan.

Les résultats de l'analyse hiérarchique ascendant sont présentés sous forme d'un dendrogramme [20-21]. Average linkage hiérarchique agglomération avec la distance Euclidienne est adopté pour les données spectrales infrarouges

Dans cette étude, PLS-DA a été appliquée en utilisant l'algorithme PLS2, qui laisse placer une matrice de \mathbf{Y} avec plusieurs variables, une par chacune des classes représentées, d'Ait Baha, de Guelmim et d'Agadir. Pour chaque classe, la variable y est placée égale à 1 pour les échantillons appartenant à la classe et à 0 pour ceux qui n'appartiennent pas à la classe considérée.

La taille du modèle de PLS-DA a été choisie à l'aide de la validation croisée. Le nombre de variables latentes a été choisi comme meilleur pour une faible valeur de RMSEP et un nombre décroissant d'échantillons mal classés. Le modèle a été finalement validé sur un ensemble externe de neuf échantillons de validation.

Le paquet d'UNSCRAMBLER 10 (CAMO, Norvège) a été employé pour tous les prétraitements et analyse statistique de données.

3. Résultats et discussions

3.1. Interprétation des prétraitements des données spectrales

Les spectres moyen infrarouge des 30 échantillons d'amandes d'argan des trois régions du Maroc ont été enregistrés et divisés en deux groupes.

Pour cette étude exploratoire, 21 échantillons pour la calibration et 9 échantillons pour la prédiction (validation externe). La Figure 10 montre les spectres FTIR des amandes d'argan étudié. Les différences entre eux sont clairement petites et produites seulement dans des régions limitées des spectres. L'utilisation de pics caractéristiques ou des plages de longueur d'onde étroites pour obtenir des informations utiles pour distinguer les amandes semblait très difficile.

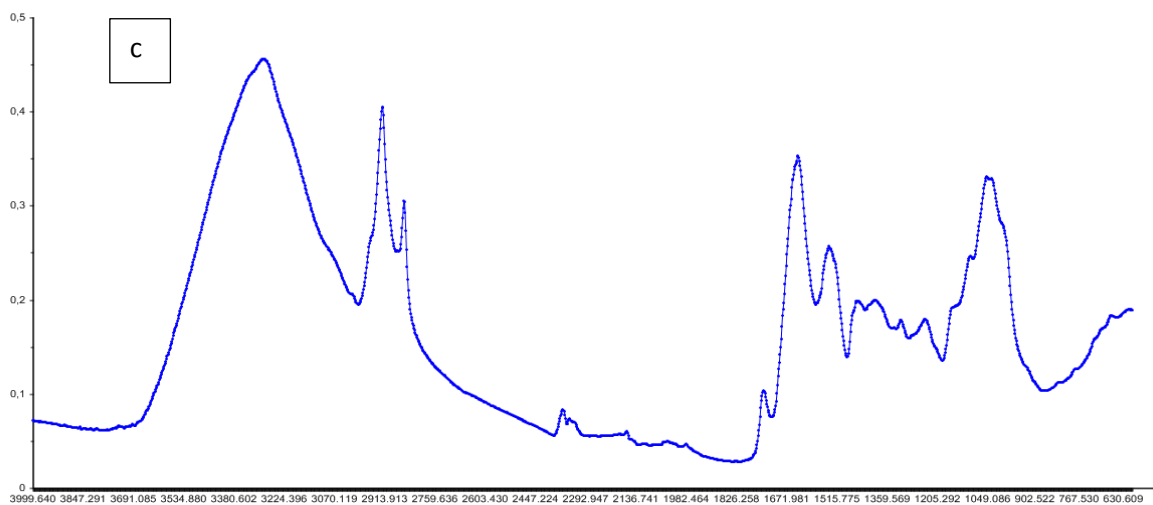
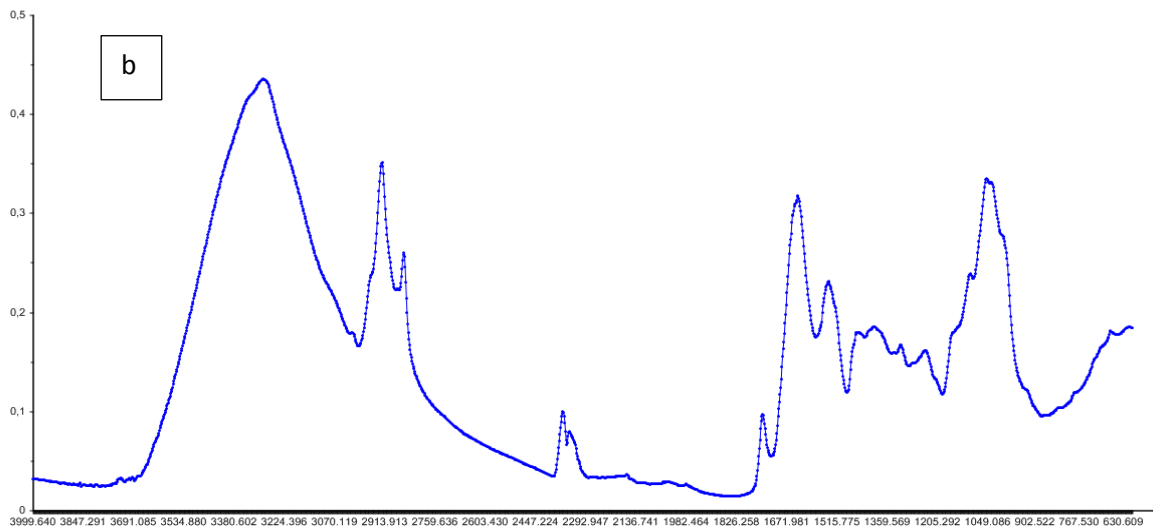
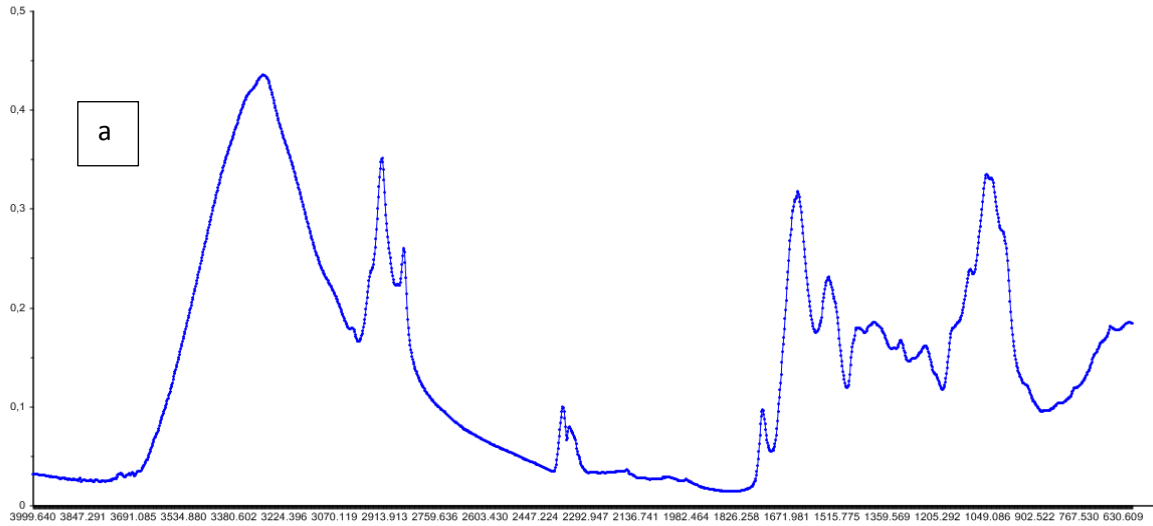


Figure 10 : Spectre moyen infrarouge des amandes d’argan issues des régions d’Agadir (a), d’Ait baha(b) et de Guelmim(c) respectivement

La bande de l'eau au voisinage de 2994 cm^{-1} n'est pas supprimée pour cette étude. Les pics autour de $2800\text{-}3080\text{ cm}^{-1}$ sont caractéristiques des vibrations d'élongation du groupe C-H, un autre pic entre $1700\text{ et }1800\text{ cm}^{-1}$ est attribuable à la bande d'élongation du groupe C=O. Les bandes d'élongations du C-O-C et la bande C-H de flexion dans la région $900\text{-}1400\text{ cm}^{-1}$ pourraient être facilement observées [23].

Dans le but d'obtenir une information plus claire des données spectrales de FTIR, les spectres ont subi premièrement des prétraitements mathématiques. En particulier, déviation normale standardisée (SNV) et les prétraitements de dérivés ont été examinés [24-25].

Les meilleurs résultats ont été obtenus avec la deuxième dérivée de Savitzky-Golay [26] avec un polynôme du deuxième ordre et lissage de 5-points. Ce prétraitement a fourni les meilleurs résultats pour PLS-DA.

3.2. La classification hiérarchique ascendante (CHA):

La classification hiérarchique ascendante permet de regrouper des individus ou des variables par ordre de proximité croissante. La méthode de mesure des distances utilisée a été celle des distances euclidiennes représentant les distances géométriques à l'intérieur de l'espace multidimensionnel [27].

Les méthodes de classification ont pour objectif de regrouper des objets en un nombre restreint de classes homogènes [28].

HCA est une technique non supervisée qui emploie l'information obtenue à partir des variables mesurées pour indiquer les similitudes ou la dissimilitude existant entre les échantillons étudiés [29-30].

Les résultats de l'analyse hiérarchique ascendante ont été obtenus avec la deuxième dérivée de type Savitzky-Golay avec un polynôme du deuxième ordre et un lissage de 5-point.

La figure 11 montre le dendrogramme obtenu à partir de la classification hiérarchique ascendante (CHA) en appliquant la distance euclidienne.

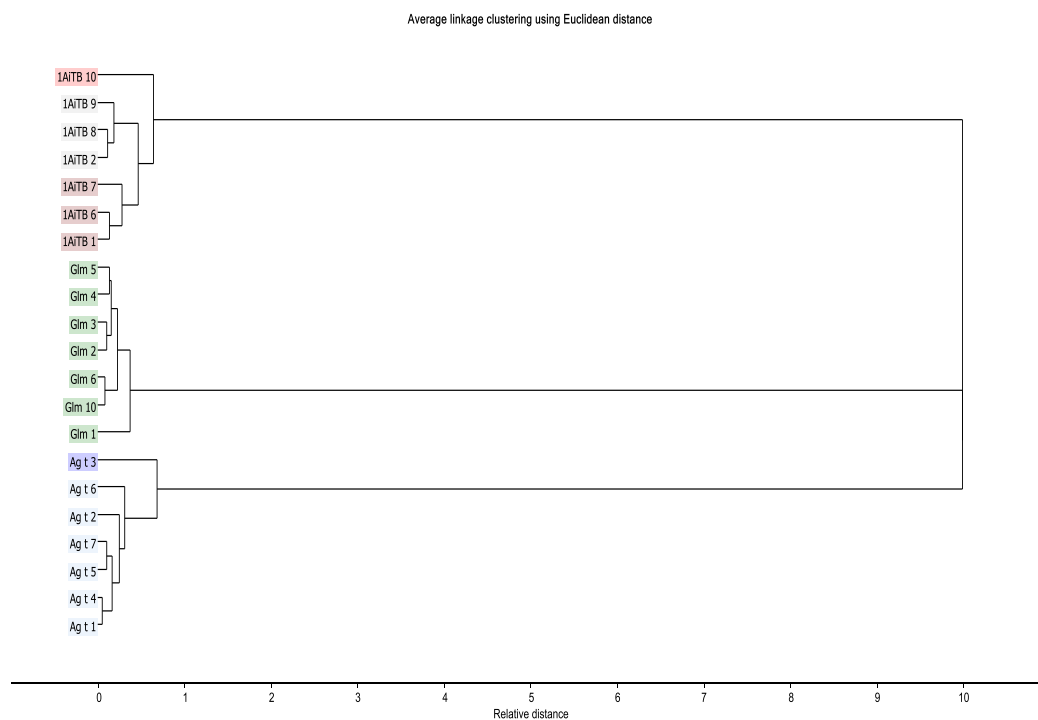


Figure 11 : Dendrogramme avec la distance euclidienne pour les trois groupes des amandes d’argan

Trois groupes peuvent être clairement identifiés. Le premier groupe contient sept échantillons rassemblés dans le groupe d’Ait Baha. Le deuxième se compose de sept échantillons, tous localisés dans Guelmim, Le dernier groupe se compose de sept échantillons, tous sélectionnés dans le territoire d’Agadir. Le dendrogramme prouve que les spectres IRTF tiennent compte d’une définition très précise de l’origine des trois types d’amandes d’argan. Comme il peut être vu dans le dendrogramme, les trois origines mènent à rendre très bien des faisceaux compacts, formés toujours par des échantillons provenant de la même origine. Ces trois groupes sont clairement séparés l’un par rapport à l’autre.

3.3 Classification PLS-DA et prédiction

Dans l’étape de la classification, PLS-DA a été exécuté dans le but d’établir un modèle de classification utile pour la prédiction postérieure de nouveaux échantillons des amandes d’argan d’origine inconnue. La procédure a été basée sur un algorithme PLS2, en fixant la valeur de la variable Y (origine géographiques) et X les spectres MIR pour chacune des trois classes : c’est 1 pour les échantillons appartenant à la classe et 0 pour le reste, c.-à-d., 1 est attribué pour des échantillons dans la classe et 0 pour des échantillons hors de la classe. Différents prétraitements ont été examinés, comme la variable aléatoire normale standard

(SNV) et la 1^{ère} et 2^{ème} dérivées avec différents ordres et tailles polynômes. Pour ces traitements, le centrage et la réduction des variables ont été également examinés.

Les résultats obtenus par la PLS-DA sont récapitulés dans le tableau 9. Pour chacun des modèles établis, les paramètres globaux tels que le nombre de variables latentes dans le modèle et la variance expliquée de Y. Il est impossible de savoir *a priori* lequel des prétraitements donneront les meilleurs résultats. Il faut donc tester de manière systématique les différents prétraitements et sélectionner celui donnant la RMSEP la plus faible.

Pour chacune des trois classes utilisées dans le modèle : Ait Baha, (- aitb), Agadir (- agt) et Guelmim (- glm), le RMSEP et le coefficient de corrélation carré entre **yprédite** et **ymesurée** sont donnés. En outre, la qualité de classification pour chaque classe dans chaque modèle est évaluée en fournissant le nombre de positifs faux (échantillons identifiés comme dans la classe, alors qu'ils ne les sont pas) et de négatifs faux (échantillons identifiés comme hors de la classe alors qu'ils les sont). À cette fin, une valeur seuil de 0.5 a été choisie et des échantillons avec une valeur prévue de y au-dessus de 0.5 sont considérés dans la classe et en-dessous de 0.5 comme hors de la classe.

Tableau 9 : Les paramètres statistiques obtenus avec la validation croisée pour le modèle PLS-DA

| Prétraitement(*) | Classe | Nombre de variable | Y variance axpliqué (%) | RMSEP | R ² V | False + | False - |
|--------------------|--------|--------------------|-------------------------|--------|------------------|---------|---------|
| raw data (none) | -aitb | 3 | 47.06 | 0.2524 | 0.7566 | 2 | 0 |
| | -agt | | | 0.4875 | 0.0215 | 2 | 3 |
| | -glm | | | 0.3236 | 0.6057 | 1 | 1 |
| SNV CN | -aitb | 6 | 63.54 | 0.2412 | 0.7672 | 0 | 1 |
| | -agt | | | 0.4098 | 0.3469 | 3 | 2 |
| | -glm | | | 0.2366 | 0.7859 | 0 | 0 |
| 2D2O9S C | -aitb | 7 | 75.47 | 0.1660 | 0.8850 | 0 | 0 |
| | -agt | | | 0.2990 | 0.6351 | 1 | 0 |
| | -glm | | | 0.2367 | 0.7527 | 1 | 1 |
| 2D2O9SCN | -aitb | 6 | 69.29 | 0.1820 | 0.8504 | 0 | 0 |
| | -agt | | | 0.3493 | 0.4818 | 2 | 2 |
| | -glm | | | 0.2411 | 0.7485 | 0 | 0 |
| 1D2O5S C | -aitb | 7 | 64.39 | 0.2110 | 0.8129 | 0 | 0 |
| | -agt | | | 0.3807 | 0.4333 | 1 | 2 |
| | -glm | | | 0.2694 | 0.6977 | 1 | 1 |
| 1D2O5S CN | -aitb | 6 | 80.73 | 0.1675 | 0.8836 | 0 | 0 |
| | -agt | | | 0.2850 | 0.6727 | 1 | 2 |
| | -glm | | | 0.1881 | 0.8626 | 0 | 0 |

Y a expliqué le désaccord, R²V : coefficient de corrélation de validation.

(*) C : Centrage moyen et CN centré normé. Pour le prétraitement dérivé, par exemple, 1DO5S, 1D ou 2D dérivé de moyen 1^{ère} ou 2^{ème}, respectivement, O signifie le polynôme du deuxième ordre et le 5S cinq points de taille de fenêtre.

La classe d'Agadir est indépendante du prétraitement utilisé, il est toujours la plus difficile à modeler, ce qui peut être clairement vu par le RMSEP, et la valeur de R^2 .

Le tableau 9 montre que la dérivée première avec le polynôme du second degré et la fenêtre de cinq points fournissent les meilleurs résultats en termes de nombre d'erreurs de classification, RMSEP et R^2 CV et une forte variance Y expliquée. D'autres options, telles que la dérivée seconde avec un polynôme de second ordre et une fenêtre de 7 ou 9 points avec centrage sont aussi des alternatives appropriées en termes de taux de classification, mais semblent être légèrement pire pour les RMSEP et R^2 CV observées.

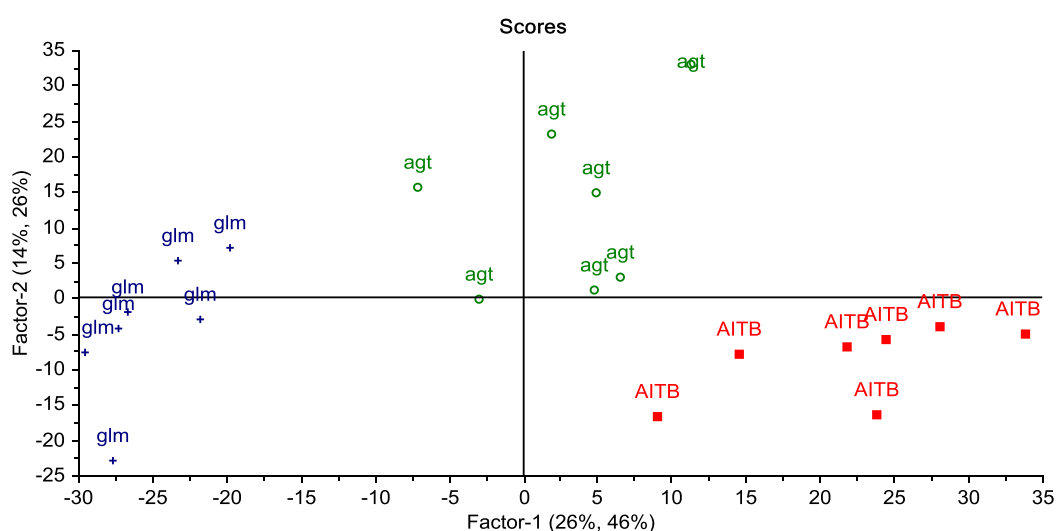
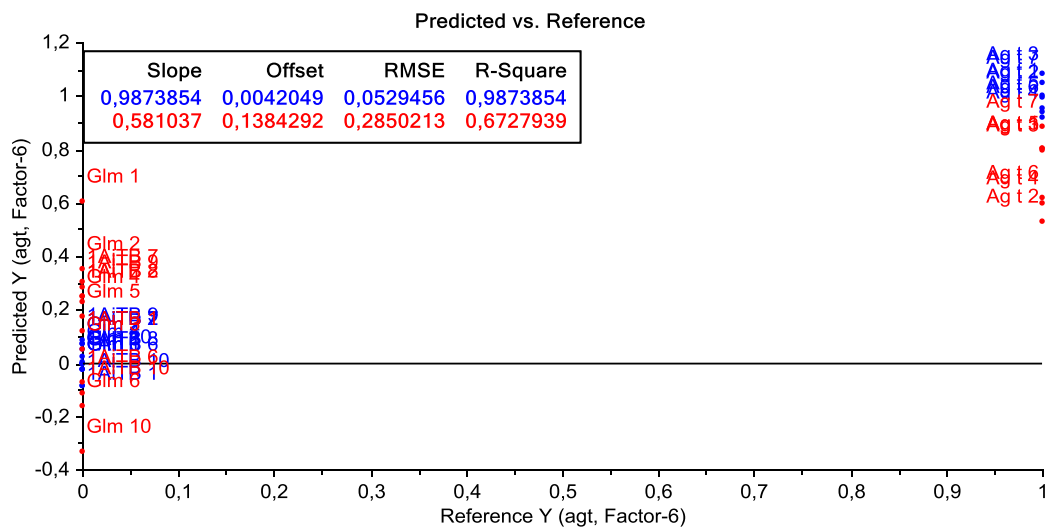
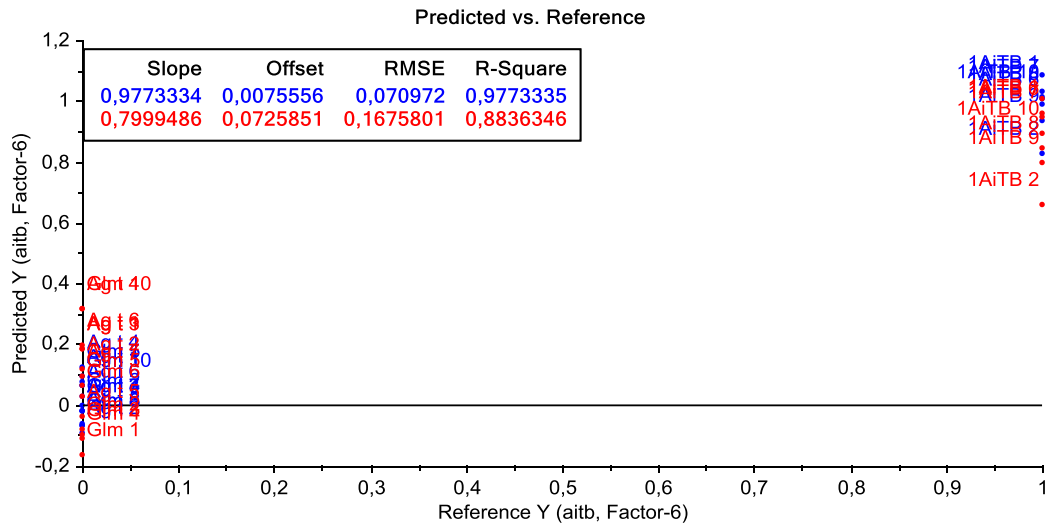


Figure 12: Score plot pour le modèle PLS-DA des amandes d'argan pour les trois régions Ait baha (AITB), Agadir (agt) and Guelmim(glm) avec le prétraitement premier dérivée

En se concentrant sur le modèle PLS-DA finalement retenu, la figure 12 montre le score plot PLS-DA pour les échantillons d'étalonnage. Avec l'utilisation de deux variables latentes, 72% de la variance Y est déjà expliquée. Trois régions (groupe) peuvent être clairement identifiées, chacun d'eux représentant un groupe des amandes d'argan d'une région. Ce résultat montre clairement la capacité potentielle de cette technique de discriminer entre les trois origines des amandes d'argan utilisées dans cette étude. Dans le même temps, il convient de noter que la classe Agadir est placée entre les deux autres.

La figure 13 montre les valeurs de y prédites et la valeur y mesurée pour chaque classe. Les valeurs en rouge donnent les résultats de validation croisée obtenus pour le modèle de calibration (tableau 9)



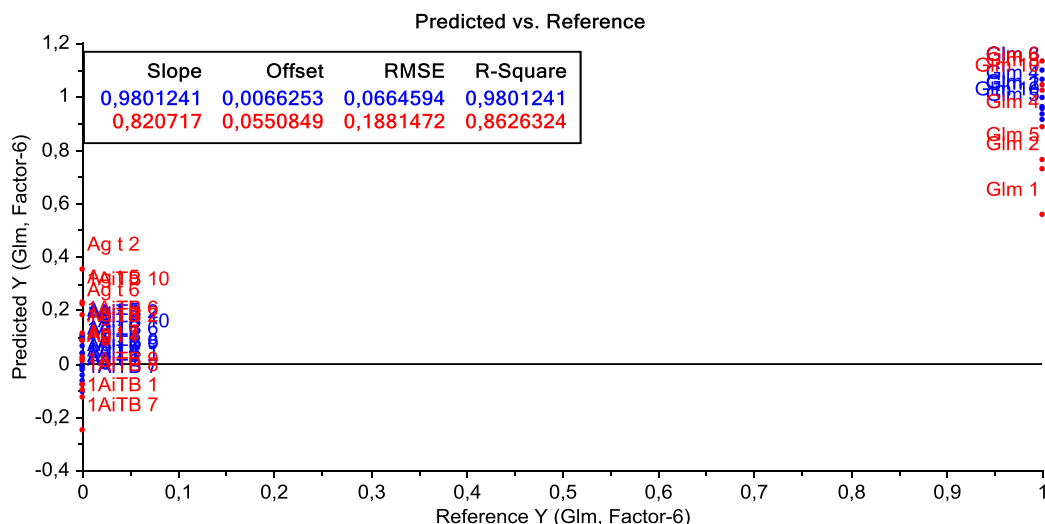


Figure 13: Les valeurs prédites en fonction des valeurs mesurées pour Ait Baha, Agadir et Guelmim avec le modèle PLS –DA

La classe Ait Baha donne les meilleurs résultats puisque les valeurs de la calibration et de validation croisée sont très similaires. La classe Guelmim donne également des résultats bien modélisés pour la validation croisée, bien que les valeurs de y prédites sont plus en propagation (moins compact en «in and out» des groupes de classe) que les résultats de l'étalonnage. La grande différence est notée lorsqu'on compare les résultats d'étalonnage et de validation croisée de la classe d'Agadir. Néanmoins, le taux de classement est encore satisfaisant, bien que certains échantillons restent très proches de la limite de séparation des classes.

Dans les trois classes, Agadir et Guelmim sont principalement responsables de false négative du 2 échantillon d'Agadir un false négatif pour la classe Guelmim car la classe d'Agadir est le plus difficile à classer par rapport aux autres classes.

Classification des nouveaux échantillons (prédiction)

Le modèle de classification choisi a été appliqué sur le groupe de validation externe FTIR, formé par neuf nouveaux échantillons des amandes d'argan, trois pour chaque classe. Cette étape a permis de tester la capacité du modèle construit pour classer les nouveaux échantillons dans les catégories précédemment établies.

Le tableau 10 montre les résultats de classification avec la comparaison entre les valeurs prédites et les valeurs de références attendues pour chaque échantillon pour chaque variable de classe.

Tableau 10: Les résultats de classification des échantillons de validation externe des amandes d'argan.

| | AIT BAHA | | AGADIR | | GUELMIM | |
|--------------|---------------|-----------|---------------|-----------|---------------|-----------|
| | Prédiction | Référence | Prédiction | Référence | Prédiction | Référence |
| Ait b | 0.9490 | 1 | -0.4027 | 0 | 0.4538 | 0 |
| Aitb | 1.0109 | 1 | -0.1153 | 0 | 0.1044 | 0 |
| Ait b | 1.0748 | 1 | -0.1558 | 0 | 0.0810 | 0 |
| Agt | 0.0748 | 0 | 0.9681 | 1 | -0.0445 | 0 |
| Agt | 0.0812 | 0 | 0.5924 | 1 | 0.3264 | 0 |
| Agt | 0.0324 | 0 | 0.9914 | 1 | -0.0239 | 0 |
| Glm | 0.0050 | 0 | 0.1001 | 0 | 0.8949 | 1 |
| Glm | 0.0264 | 0 | -0.1748 | 0 | 1.1484 | 1 |
| Glm | 0.0342 | 0 | 0.1442 | 0 | 1.1100 | 1 |

Les valeurs en gras indiquent les échantillons qui confirment l'appartenance à la classe prévue. D'après Le tableau les valeurs de Y prédites, sont proches de 1 et les neufs échantillons sont classés dans les trois groupes. Ces résultats confirment la robustesse du modèle PLS-DA développée à base des spectres MIR.

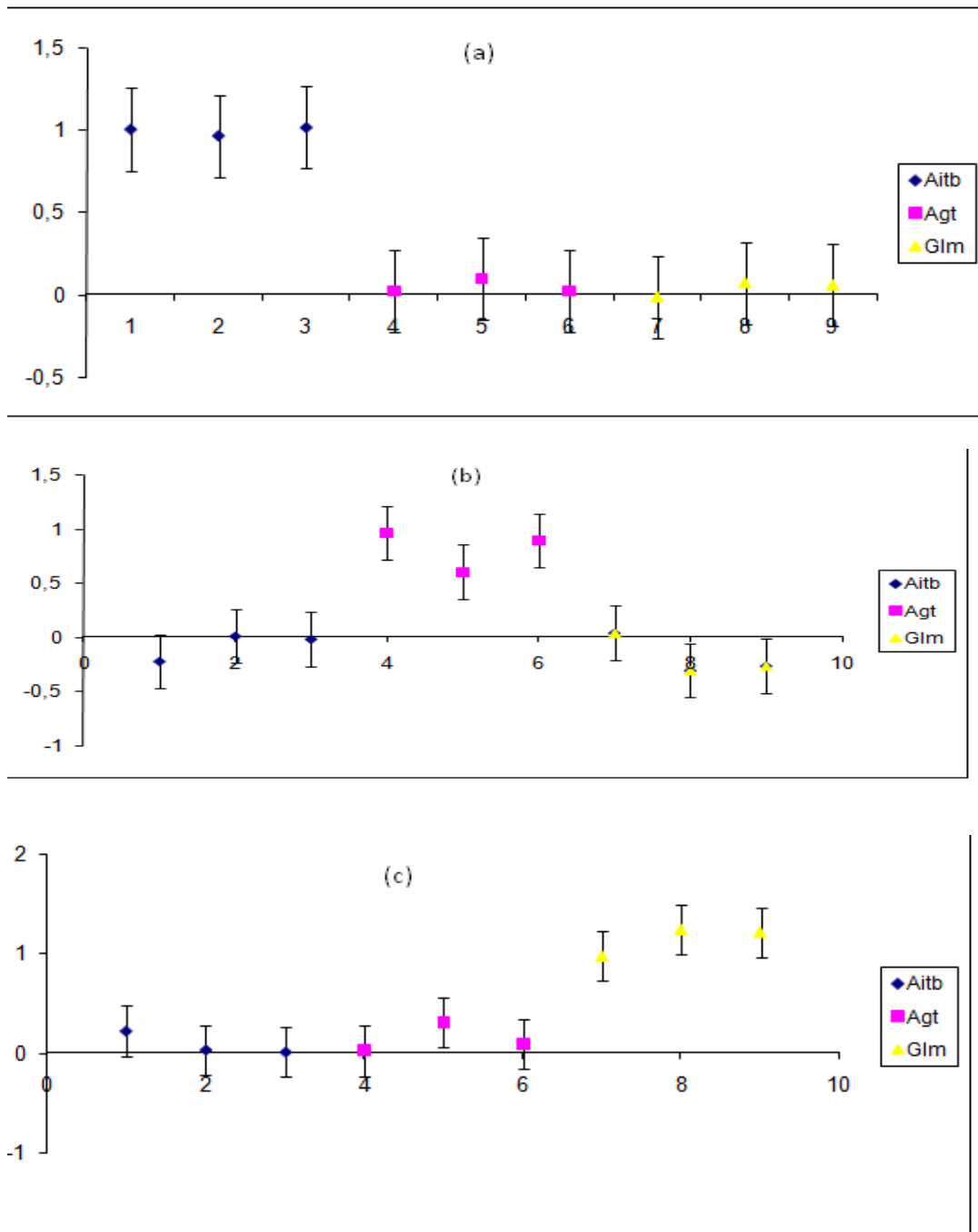


Figure 14 : La Prédictions de validation externe des amandes d’argan : (a) classe Ait Baha; (b) classe Agadir et (c) la classe Guelmim.

Les figures 14(a), (b) et (c) illustrent plus en détail les résultats du tableau 10, en plus de la barre d’erreur de prédiction pour chaque échantillons du classe. La figure 14a montre les résultats de la classification pour la variable liée à la classe d’Ait Baha. C’est la classe la mieux définie et les échantillons dans la classe sont bien au-dessus de la valeur seuil de 0,5 même en considérant l’erreur associée à la prédiction, et des échantillons sur de la classe sont

très bien reconnu aussi. La figure 14c montre une affectation correcte des échantillons de Guelmim dans son origine. On peut dire que les échantillons de la classe sont classés correctement. Dans la figure 14b, les résultats de la classe d'Agadir sont représentés. Bien que les échantillons d'Agadir se classent dans la classe, l'échantillon 5 est classé hors des limites lorsque l'erreur associée à la prédiction est considérée. Cela confirme la difficulté liée à la bonne modélisation de cette classe.

Conclusion

L'origine des échantillons des amandes d'argan peut être distinguée en combinant la mesure de spectres IR et les outils chimiométriques pour l'analyse discriminante. Une procédure d'analyse multivariée basée sur l'utilisation combinée de la classification hiérarchique (HCA) et Partial Least Squares-Analyse Discriminante (PLS-DA) a été testée à cet effet.

L'analyse en composantes hiérarchiques HCA et la classification PLS-DA ont été utilisées afin de construire un modèle de classification pour vérifier l'origine des amandes d'argan à partir de mesures FTIR. L'utilisation combinée de cette technique d'analyse et des méthodes chimiométriques adéquates s'est avéré être une stratégie efficace pour discriminer les amandes d'argan en fonction de leur origine géographique sans avoir besoin d'autres méthodes d'analyse chimique. HCA a montré trois groupes distincts liés aux régions marocaines et a suggéré l'utilité de la spectroscopie infrarouge à distinguer entre les origines du noyau. Les trois régions étudiées pourraient être clairement classées, bien que la région d'Agadir a été plus difficile à modéliser. Cette méthode peut être proposée comme une alternative appropriée pour le contrôle de la qualité des amandes d'argan et la détermination de leur origine géographique, cette stratégie d'authentification peut être effectuée avant la fabrication de l'huile d'argan finale.

Références bibliographiques

- [1] UNESCO 2007. “Biosphere Reserve Information” MAB Programme.
- [2] Cherki, M., Berrougui, H., Drissi, A., Adlouni, A., & Khalil, A. 2006. Argan oil: Which benefits on cardiovascular diseases? *Pharmacological Research*, 54, 1–5.
- [3] Downey, G., McIntyre, P., & Davies, A. N. 2003. Geographic classification of extra virgin olive oils from the eastern Mediterranean by chemometric analysis of visible and near-infrared spectroscopic data. *Applied Spectroscopy*, 57(2), 158–163.
- [4] Marcos-Lorenzo, I., Pérez-Pavón, J. L., Fernández-Laespada, M. E., García-Pinto, C., & Moreno-Cordeo, B. 2002. Detection of adulterants in olive oil by headspace mass spectrometry. *Journal of Chromatography A*, 945, 221–230.
- [5] Sayazo, A., Morales, M. T., & Aparicio, R. 2004. Detection of hazelnut oil in virgin olive oil by a spectrofluorimetric method. *European Food Research and Technology*, 218, 480–483.
- [6] Reid, L. M., O'Donnell, C. P., & Downey, G. 2006. Recent technological advances for the determination of food authenticity. *Trends in Food Science and Technology*, 17(7), 344–353. doi:10.1016/j.tifs.2006.01.006.
- [7] Aparicio, R.; Aparicio-Ruiz, R. 2000. Authentication of vegetable oils by chromatographic techniques. *J. Chromatogr*, 881, 93-104.
- [8] Spangenberg, C. E.; Macko, S. A.; Hunziker, J. 1998. Characterization of olive oil by carbon isotope analysis of individual fatty acids: Implications for authentication. *J. Agric. Food Chem.*, 46, 4179-4184.
- [9] Stefanoudaki, E.; Kotsifaki, F.; Koutsaftakis, A. 1999. Classification of virgin olive oils of the two major Cretan cultivars based on their fatty acid composition. *J. Am. Oil Chem*, 76, 623-626
- [10] Tsimidou, M. Papadopoulos, G. et Boskou, D. 1992. Determination of phenolic compounds in virgin olive oil by reverse-phase HPLC with emphasis on UV detection. *Food Chemistry*, 44 ,53-60.
- [11] Alberghina, G.; Caruso, L.; Fisichella, S.; Musumarra, G. 1991. Geographical classification of Sicilian olive oils in terms of sterols and fatty-acids content. *J. Sci. Food Agric.*, 56, 445-455.

- [12] Tasioula-Margari, M.; Okogeri, O. 2001. Isolation and characterization of virgin olive oil phenolic compounds by HPLC/UV and GCMS. *J. Food Sci.*, 66, 530-534.
- [13] Garcia, A.; Brenes, M.; Romero, C.; Garcia, P.; Garrido, A. 2002. Study of phenolic compounds in virgin olive oils of the Picual variety. *Eur. Food Res. Technol.*, 215, 407-412
- [14] Webster, L.; Simpson, P.; Shanks, A. M.; Moffat, C. F. 1999. The authentication of olive oil on the basis of hydrocarbon concentration and composition. *Analyst*, 125, 97-104
- [15] Bortolomeazzi, R.; Berno, P.; Pizzale, L.; Conte, L. S. 2001. Sesquiterpene, alkene, and alkane hydrocarbons in virgin olive oils of different varieties and geographical origins. *J. Agric. Food Chem.*, 49, 3278-3283.
- [16] Hilali, M.; Charrouf, Z.; Soulhi, A.; Hachimi, L.; Guillaume, D. 2005. Influence of Origin and Extraction Method on Argan Oil Physico-Chemical Characteristics and Composition. *J. Agric. Food Chem.*, 53, 2081-2087
- [17] Lai, Y. W., Kemsley, E. K., & Wilson, R. H. 1994. Potential of Fourier transform infrared spectroscopy for the authentication of vegetable oils. *Journal of Agricultural and Food Chemistry*, 42, 1154–1159.
- [18] Dahlberg, D. B., Lee, S. M., Wenger, S. J., & Vargo, J. A. 1997. Classification of vegetable oils by FT-IR. *Applied Spectroscopy*, 51, 1118–1124.
- [19] Dupuy, N., Duponchel, L., Huvenne, J. P., Sombret, B., & Legrand, P. 1996. Classification of edible fats and oils by principal component analysis of Fourier transform infrared spectra. *Food Chemistry*, 57, 245–251.
- [20] M'hirit, O., Benzyane, M., Benchekroun, F., El yousfi, M., Bendaanoun 1998. L'arganier, une espèce fruitière à usage multiples, Ouvrage collectif Edition Mardaga (Belgique), 150.
- [21] Jolliffe, T., O.B. Allen, B.R. Christie, 1989. Comparison of variety means using cluster analysis and dendrograms, *Experimental Agriculture*, 25, 259-269.
- [22] Massart, D.L., Kaufman, L. 1983. The interpretation of analytical chemical data by the use of cluster analysis. New York, Wiley.
- [23] Tay, A., Singh, R. K., Krishnan, S. S., & Gore, J. P. 2002. Authentication of olive oil adulterated with vegetable oils using Fourier transform infrared spectroscopy. *LWT-Food Science and Technology*, 35, 99–103.

- [24] Iñón, F. A., Garrigues, J. M., Garrigues, S., Molina, A., & de la Guardia, M. 2003. Selection of calibration set samples in determination of olive oil acidity by partial least squares-attenuated total reflectance-Fourier transform infrared spectroscopy. *Analytica Chimica Acta*, 489, 59–75.
- [25] Maggio, R. M., Kaufman, T. S., Del Carlo, M., Cerretani, L., Bendini, A., Cichelli, A., et al. 2009. Monitoring of fatty acid composition in virgin olive oil by Fourier transformed infrared spectroscopy coupled with partial least squares. *Food Chemistry*, 114, 1549–1554.
- [26] Savitzky, A.; Golay, M.J.E. 1964. "Smoothing and Differentiation of Data by Simplified Least Squares Procedures". *Analytical Chemistry* 36 (8), 1627–1639
- [27] Falissard, B. 1996. *Comprendre et utiliser les statistiques dans les sciences du vivant*, Masson eds, 314, Paris.
- [28] Lebart, L., Morineau, A. & Tabard, N. 1977. *Technique de la description statistique, méthodes et logiciels pour l'analyse des grands tableaux*. Paris: Dunod.
- [29] Peterson, L.E., 2002. Software Report: CLUSFAVOR 5.0: hierarchical cluster and principal component analysis of microarray-based transcriptional profiles, *Genome Biol*, 3(7): 1-8.
- [30] Skrobot, V.L. E.V.R. Castro, R.C. Pereira, V.D. Pasa, I.P. Fortes, 2005. Identification of Adulteration of Gasoline Applying Multivariate Data Techniques HCA and KNN in Chromatographic Data, *Energ. Fuel.*, 19(6), 2350-2356.

Chapitre 6 : CARACTERISATION PAR CHROMATOGRAPHIE EN PHASE GAZEUSE DES ACIDES GRAS DES HUILES

1. Introduction

La demande des huiles végétales de haute qualité est d'une grande importance dans un régime parfaitement équilibré. Par conséquent, il existe un besoin pour des techniques d'analyse afin de contrôler la pureté et l'authenticité des huiles. Les méthodes de contrôle qualité des acides gras (AG) sont la détermination de l'indice de peroxyde, l'indice d'iode, l'indice de réfraction, le temps Rancimate d'induction et trans-valeur [1-3]. Bien que ces méthodes soient utiles avec des échantillons purs, la diversité des méthodes de falsification exigent des techniques plus sophistiquées.

Dernièrement, diverses techniques ont été proposées pour établir l'authenticité des huiles végétales et pour détecter la falsification telles que la spectrophotométrie UV [4], la Spectroscopie Raman [5,6], la RMN ¹³C [7] et la spectrométrie IR [8,9]. Ces méthodes ont été utilisées pendant de nombreuses années dans l'analyse de l'huile et l'industrie des graisses.

Cependant, ces techniques présentent des limites en tant que des méthodes de détection du fait que les différences spectrales de la plupart des huiles végétales sont tout à fait petites (contiennent la même AF).

Cependant, chaque huile végétale a sa propre composition en acide gras [10]. Par conséquent, la détermination de l'acide gras "empreinte digitale" par des méthodes chromatographiques fournit des informations utiles concernant l'authenticité et la possibilité de la falsification des huiles. Généralement, la détermination des acides gras est réalisée par Chromatographie liquide (LC) [11,12] et de la chromatographie en phase gazeuse (CPG) [13,14].

2. Matériels et méthodes

2.1. Acides gras de référence

Deux acides gras saturés de pureté comprise entre 97% et 99,5% (Aldrich) et deux acides gras insaturés de puretés égales à 99% (Sigma) ont été choisis pour établir une référence spectrale des lipides. Le code correspondant à chaque acide gras est composé d'un chiffre indiquant le nombre de doubles liaisons, d'une lettre et du nombre d'atomes de carbone dans la chaîne de l'acide gras.

Pour les acides gras saturés, le premier chiffre est « 0 », la lettre est <C> pour « carbone » et le nombre d'atomes de carbone est compris entre 6 et 20.

Concernant les acides gras insaturés :

Le premier chiffre est « 1 » ou « 2 » respectivement pour une ou deux doubles liaisons

Le code, le pourcentage de pureté, le fabricant et les points de fusion des acides gras étudiés sont donnés dans le tableau 11.

Tableau 11 : Acides gras estérifiés, saturés et insaturés (points de fusion tiré de : Handbook of Chem. & Phys.1968/69)

| Esters des acides gras saturés | Formule | Firme | Pureté | CODE | Point de fusion |
|--|--------------------|-------------|--------|------|-----------------|
| acide palmitique (acide hexadécanoïque) | C ₁₆ :0 | Aldrich USA | 99% | 0C16 | 30°C |
| acide stéarique (acide octadécanoïque) | C ₁₈ :1 | Aldrich USA | 97% | 0C18 | 39,1°C |
| Acides gras insaturés | | | | | |
| acide oléique (acide octadécénoïque) | C18:1 <i>cis</i> | Sigma, D. | 99% | 1C18 | 5°C |
| acide linoléique (acide octadéca diénoïque) | C18:2w6 <i>cis</i> | Sigma, D. | 99% | 2C18 | -5°C |

2.2.Principe : estérification par le méthanol en présence d'un acide fort.

Les acides gras sont les constituants essentiels des lipides (matières grasses).

La méthode d'analyse utilisée est passée par trois étapes

1. extraction des lipides (éther, hexane,..) ;
2. saponification (OH-) pour obtenir les acides gras ;
3. transformation (dérivatisation) des acides gras en esters méthyliques

Les huiles sont principalement composées de triglycérides - triesters d'acides gras du glycérol. Ces molécules sont difficiles à analyser tel quel en chromatographie gazeuse, et on analyse généralement les acides. Pour se faire, il faut d'abord détruire les liaisons ester

glycérol - acide gras et à synthétiser des esters du méthanol et des acides gras (ces esters, étant plus volatils, sont plus faciles à analyser en CPG que les acides gras « libres »).

2.3 Protocole expérimental utilisé pour la préparation des esters méthyliques d'acides gras

5 g d'échantillons d'acide gras sont pesés dans un ballon de 100 ml puis 50ml de méthanol sulfurique est ajouté (méthanol contenant 1% en volume d'acide sulfurique, obtenu en ajoutant lentement l'acide sulfurique dans le méthanol en agitant continuellement.)

Le mélange est porté à reflux pendant 25 minutes ensuite le chauffage est interrompu par l'ajout de 30 ml d'eau par le réfrigérant.

Après refroidissement, le contenu du ballon est transversé dans une ampoule à décanter afin de séparer la phase aqueuse de la phase organique tout en rinçant le ballon plusieurs fois en utilisant 20 ml d'eau puis d'hexane et les verser dans l'ampoule à décanter, puis agiter. Les esters méthyliques sont rassemblés dans la phase organique (phase supérieur pour l'hexane). Ensuite extraire une seconde fois par 20 ml de solvant.

La phase organique est recueillie dans une seconde ampoule et laver une deuxième fois avec 10 ml d'eau, ensuite on la filtre sur coton, la phase organique est évaporer au rotavapeur pour éliminer la presque totalité du solvant.

2.4 Analyse par chromatographie en phase gazeuse des esters méthyliques d'acides gras

Les esters obtenus, souvent appelés FAME (Fatty Acid Methyl Esters), sont séparés et quantifiés par chromatographie en phase gazeuse sur une colonne suffisamment polaire pour séparer les molécules aussi bien en fonction de leur taille que de leur degré d'insaturation (les nuages électroniques des liaisons n étant susceptibles d'interagir avec des phases stationnaires polaires).

Les paramètres de la méthode d'analyse des méthyles des triglycérides sont résumés comme suit :

Appareillage : chromatographe à phase gazeuse (*Trace GC ULTRA*) couplé à un spectromètre de masse (*Polaris Q MS* à trappe ionique)

Type d'ionisation

X EI (70 eV)

Type de solvant : CH₂Cl₂

Type de colonne : VB-5 (Methylpolysiloxane à 5% phenyl), 30 m * 0.25 mm * 0.25 µm.

Tableau 12 : Conditions d'injections

| | |
|--------------------------------|------------|
| Volume d'injection | 1 µl |
| Température d'injection | 220°C |
| Température d'interface | 300°C |
| Mode d'injection | Split |
| Gaz vecteur | Hélium |
| Débit | 1.4 ml/min |

Température de la source d'ionisation : 200°C.

La base de données utilisée : NIST MS Search

3. Résultats et discussion

3.1. Utilisation de l'analyse des acides gras

Il est d'usage de débiter une analyse de recherche d'adultération par la caractérisation de la composition en acides gras de l'huile.

Cette détermination permet de détecter des ajouts d'huiles de soja à des teneurs de l'ordre de 0 à 24 % dans l'huile d'olive. L'huile d'olive possède la particularité d'être très riche en acides gras mono insaturés tels que l'acide oléique (C18 :1 > 60 %) accompagné de trois autres acides gras principaux l'acide linoléique (C 18:2), l'acide palmitique (C16 :0) et l'acide stéarique (C18 :0) c'est pour cela qu'ils sont utilisés pour cette étude. La composition en acides gras d'une huile peut varier avec son origine et quelques autres facteurs.

3.2 Composition en acides gras des huiles étudiées

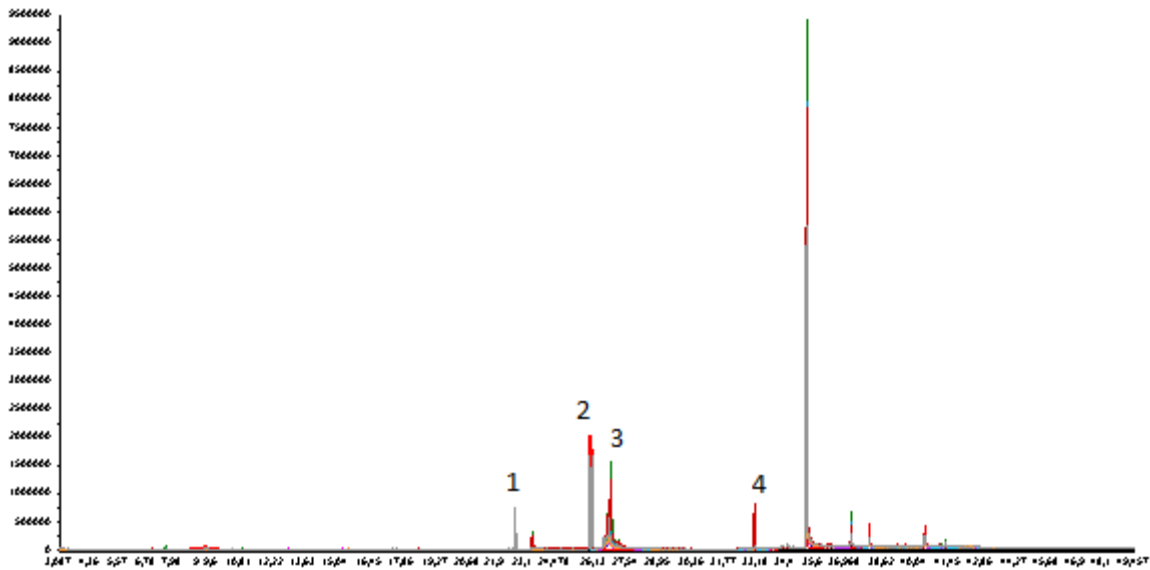


Figure 15 : Chromatogramme pour les différents mélanges de l'huile d'olive adultérée par l'huile de soja et les autres huiles végétales : 1acide palmitique, 2 acide linoléique 3, acide oléique, 4 acide stéarique

La figure présentée 15 montre une empreinte obtenue à partir de l'injection de 1ul d'esters méthyliques du mélange des huiles à analyser dans le système chromatographique.

Une première analyse visuelle permet de constater que le profil général du chromatogramme est différent. Le pic le plus intense est attribuable à l'acide oléique (C18 : 1) qui constitue 66% de l'aire totale du chromatogramme, suivi par celui de l'acide linoléique (C18 :2,) qui représente 16 %, et l'acide palmitique (C16 : 0), et l'acide stéarique (C18 : 0) et qui se distinguent respectivement par des pourcentages supérieurs à 3 et à 6%.

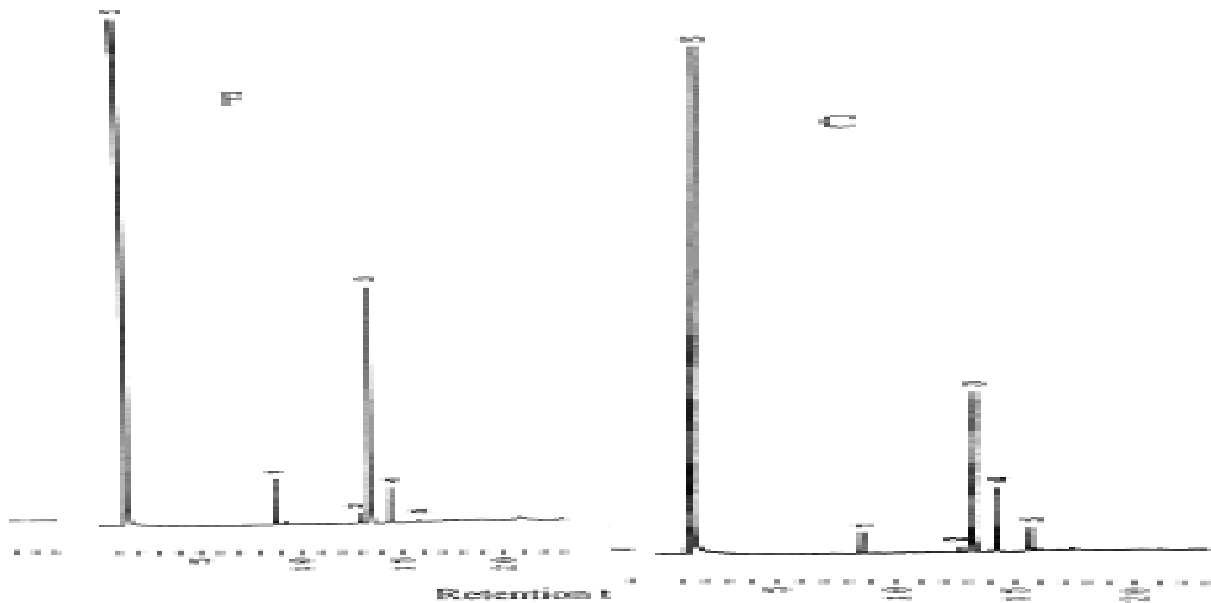


Figure 16: Chromatogramme des acides gras pour les huiles végétales : C huile de soja , F huile d'olive : 1acide palmitique, 2 acide stéarique 3, acide oléique, 4 acide linoléique, 5 acide linoléique

Tableau 13: Résultat de la composition en acides gras des mélanges d'huile d'olive avec des huiles végétales

| N° Echantillon | % d'adultération | % palmitique | % stéarique | % oléique | % linoléique |
|----------------|--------------------|--------------|-------------|-----------|--------------|
| 1 | 1.063 | 0.821 | 0.376 | 2.104 | 2.060 |
| 2 | 1.70 | 1.041 | 0.245 | 2.780 | 2.253 |
| 7 | 5.05 | 1.170 | 0 | 2.887 | 3.028 |
| 10 | 6.80 | 1.084 | 0.273 | 2.727 | 2.815 |
| 13 | 8.75 | 1.452 | 0.469 | 3.726 | 3.133 |
| 17 | 11.11 | 2.259 | 0.544 | 5.277 | 5.158 |
| 21 | 11.95 | 1.386 | 0.361 | 3.350 | 3.212 |
| T | Huile de tournesol | 5.911 | 2.241 | 23.241 | 26.762 |
| AG | Huile d'argan | 5.201 | 1.809 | 14.842 | 11.895 |
| S | Huile de soja | 6.09 | 1.85 | 16.01 | 25.01 |
| 35 | 18.03 | 1.531 | 0.637 | 4.076 | 3.511 |
| 39 | 18.47 | 3.421 | 1.064 | 1.064 | 7.104 |
| 41 | 19.23 | 7.911 | 2.356 | 20.22 | 16.574 |
| 43 | 19.79 | 5.040 | 1.681 | 12.70 | 10.35 |
| 47 | 21.44 | 3.033 | 1.143 | 8.435 | 6.985 |
| 51 | 21.64 | 1.510 | 0.511 | 2.221 | 1.806 |
| 57 | 23.00 | 6.339 | 2.136 | 16.67 | 15.31 |
| 62 | 23.41 | 6.339 | 2.136 | 16.67 | 15.31 |
| 67 | 23.84 | 5.332 | 1.709 | 12.869 | 12.375 |

Nous remarquons que la composition en acide gras de l'huile de tournesol est en accord avec les teneurs trouvées dans la littérature.

L'étude des acides gras de l'huile d'argan montre que ceux-ci sont à plus de 80% de type oléique-linoléique.

La composition pour l'huile d'argan peut présenter des variations dues à l'origine géographique du fruit de l'arganier et au climat. Une augmentation de l'acide palmitique, en passant de la plaine vers les hauts plateaux, de l'acide oléique avec la pluviométrie et de l'acide linoléique avec l'altitude ont été observées [15].

D'après le tableau 13 l'acide oléique représente le pourcentage le plus élevé pour tous les huiles sauf l'échantillon 39 suivi de l'acide linoléique et l'acide palmitique, l'acide stéarique est le plus faible pour tous les échantillons. Nous constatons à travers ces résultats, que la composition en acides gras varie selon le taux d'adultération, nous constatons que lorsque le taux d'adultération augmente le pourcentage des acides gras augmente aussi.

On peut dire qu'on ne peut pas se baser sur les ajouts d'un acide gras pour détecter l'adultération de l'huile d'olive par le soja puisque la variation du taux d'adultération engendre une variation du pourcentage de tous les acides gras.

De plus, il a été montré que d'autres facteurs affectent la composition en acide gras et spécialement la teneur en acide oléique, ce sont la variété, l'altitude, les conditions climatiques et le stade de maturation du fruit à la récolte [16-18]. Zarrouk et al, [19-20] estiment que la composition en acides gras de l'olive change en fonction des étapes de développement et de maturation. Baccouri et al, [21] ont montré que le contenu des acides gras de l'huile d'olive notamment les acides oléique, palmitique et linoléique diffère selon les variétés de l'olive étudiées et le degré de maturation du fruit.

La mise en évidence de la variation en composition en acides gras nous a conduits à réaliser une analyse en composante principale

3.3. Analyse en composantes principales

L'Analyse en composantes principales (ACP) est une méthode statistique, permettant une analyse rapide d'échantillons, sans avoir besoin d'attribuer les bandes d'absorption

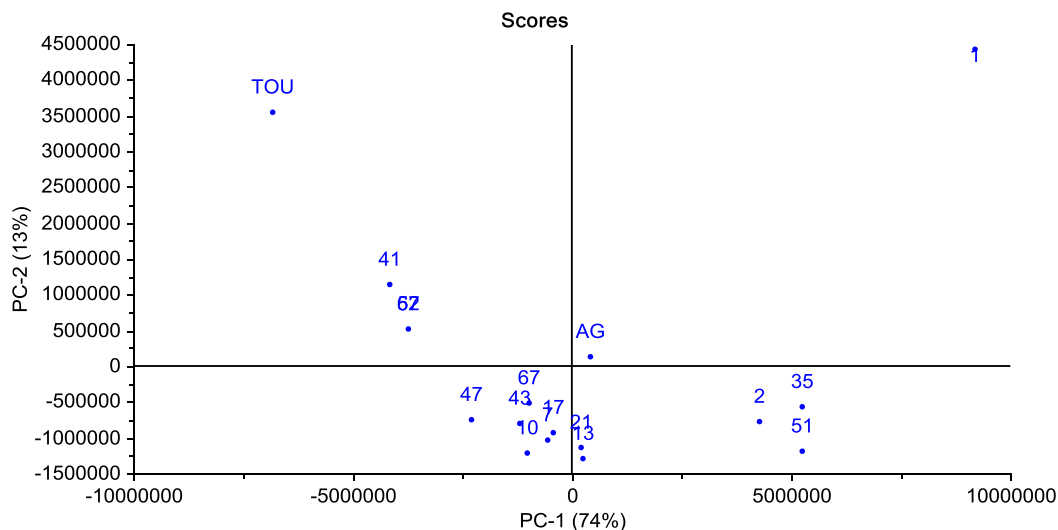


Figure 17 : Plot d'Analyse en composantes principales avec le prétraitement centré

Les données centrées sont présentées sur la figure 17. Nous remarquons que la première composante principale représente 74 % de la variance et avec deux composantes principales 88% de la variance totale est expliquée. En plus d'une séparation des échantillons selon le taux d'adultération.

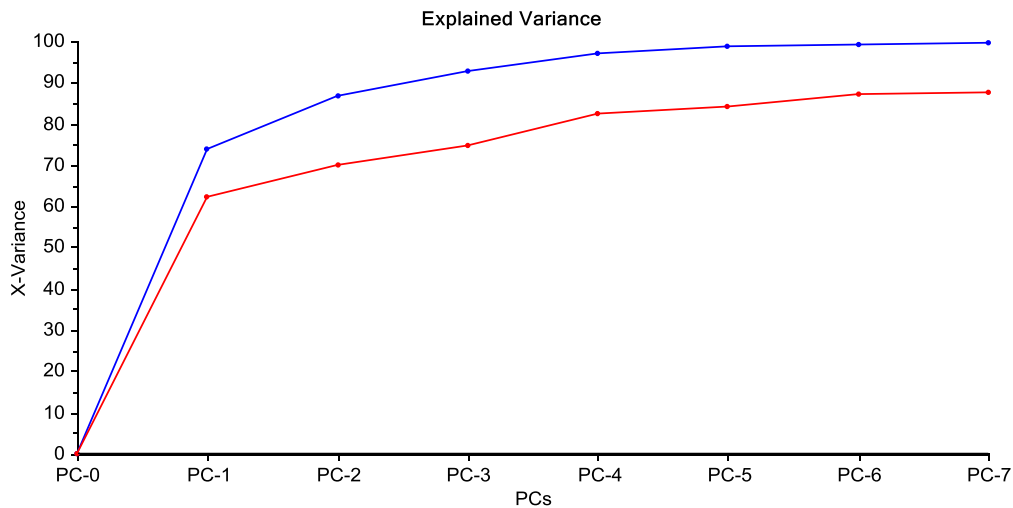


Figure 18 : La variance expliquée en fonction du nombre de composantes principales

La figure 18 présente le nombre de composantes principales en fonction de la variance

On remarque qu'avec quatre composantes principales le total de l'information est expliqué.

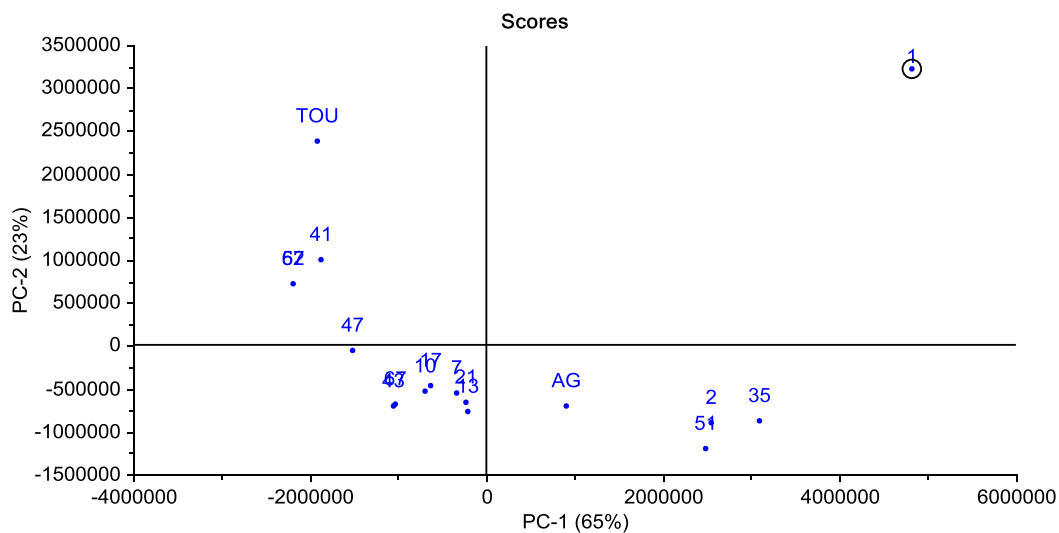


Figure 19 : Résultats de la 1^{ère} dérivée avec un polynôme égale à 2 et un smothing 3

D'après la figure 19 la première composante représente 65 % de l'information spectrale tandis que la deuxième représente 23%, les teneurs les plus faibles sont classées dans la partie négative de la 1^{ère} composante, les huiles pures sont loin des agrégats des mélanges, les teneurs élevées sont classées dans la partie positive de la 2^{ème} composante principale.

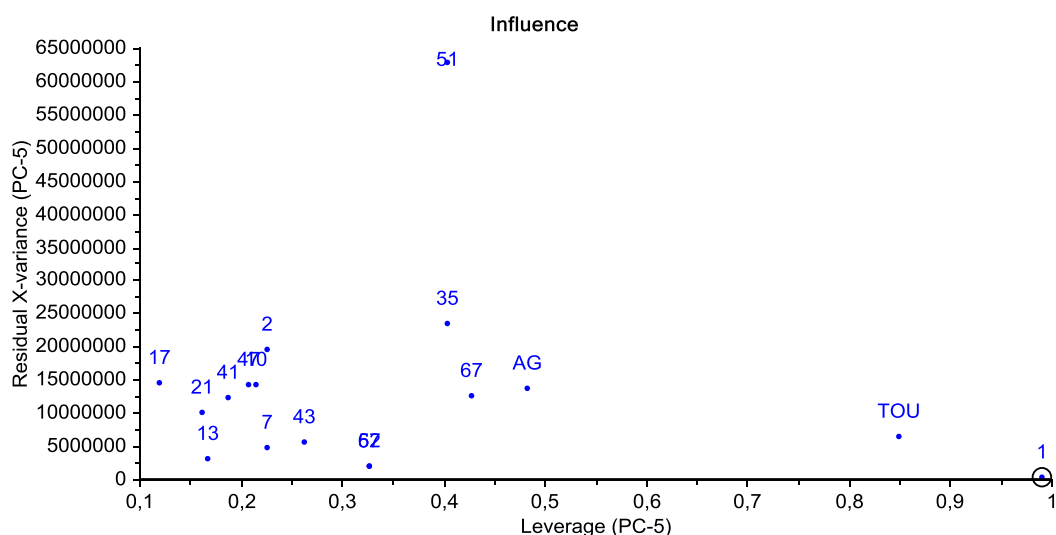


Figure 20 : Résultats de la variance résiduelle

D'après la figure 20 les huiles pures sont classées séparées des mélanges avec 5 composantes principales en plus les mélanges ayant les teneurs les plus faibles sont classées séparées des moyens teneurs et les grandes teneurs

La détermination du coefficient spécifique d'extinction dans l'ultra-violet à 270 nm (K270) et le dosage des acides gras trans, par chromatographie en phase gazeuse, peuvent être utilisés pour mettre en évidence des ajouts d'huile raffinée dans des huiles vierges. Ces critères analytiques ne sont pas suffisants pour éviter toutes fraudes [22].

L'apparition sur le marché de tournesol oléique à teneur en C18 :1 > 80 % rend cette analyse moins efficace [22].

Conclusion

Cette étude a permis de déterminer les acides gras des huiles végétales pures ainsi que les huiles adultérées. Nous n'avons pas réussi à détecter l'adultération de l'huile d'olive par l'huile de soja en se basant sur l'ajout d'un acide gras puisqu'il y a une variation de la teneur des acides gras avec la variation du taux d'adultération .

D'après des études bibliographiques la détermination de l'adultération en se basant sur la caractérisation des acides gras est un peu non informative, d'autres études sont basées sur la

caractérisation du campesterol et de la teneur en stigmastadiènes qui sont des critères très utiles [22-23].

En faisant appel à des méthodes chimiométriques comme l'analyse en composante principale les huiles végétales vierges telles que l'huile d'olive, l'huile de tournesol et l'huile d'argan sont distinguées des huiles olive adultérées par l'huile de soja.

Références bibliographiques

- [1] Harwitz (Ed.) W., 2000. Official Methods of Analysis of AOAC International, 17th ed., AOAC International, Maryland, USA, . 15 (Chapter 41).
- [2] Townshend, A. 1995. Encyclopedia of Analytical Science, vol. 3, second ed., Academic Press, London, 1540.
- [3] Labrinea, E.P. Thomaidis, N.S. Georgiou, C.A. 2001. Direct olive oil anisidine value determination by flow injection Anal. Chim. Acta, 448, 201. 201–206
- [4] Passaloglou-Emmanouillidou, S.A. 1990. A comparative study of UV spectrophotometric methods for the detection olive oil adulteration by refined oils. Zeitschrift für Lebensmittel – Untersuchung und – Forschung, Lebensm Z. Unters. Forsch., 191, 132–134
- [5] Davies A. N., McIntyre P. & Morgan E. 2000. Study of the use of molecular spectroscopy for the authentication of extra virgin olive oils. part I: Fourier transform Raman spectroscopy. Appl. Spectrosc. 54, 1864–1867
- [6] Baeten, V. Hourant, P. Morales, M.T. Aparicio, R. 1998. Oil and Fat Classification by FT-Raman Spectroscopy J. Agric. Food Chem. 46, 2638–2646.
- [7] Valhov, G. Prog. Nucl. Magn. Reson. Spectrosc. 1999, 35, 341.
- [8] Sato, T., 1997. Application of Near Infrared Spectroscopy for the Analysis of Fatty Acid Composition, Lipid Technol. 9, 46–49.
- [9] Lai, Y. W., Kemsley, E. K., & Wilson, R. H. 1994. Potential of Fourier transform infrared spectroscopy for the authentication of vegetable oils. Journal of Agricultural and Food Chemistry, 42, 1154–1159.
- [10] Spencer GF, Herb SF, Gormisky PJ 1976. Fatty acid composition as a basis for identification of commercial fats and oils. J. Am. Oil Chem. Soc. 53, 94–96.
- [11] Hein, M. Isengard, H.D. 1997. Determination of underivatized fatty acids by HPLC, Lebensm Unters. Z. Forsch. A 204, 420–424
- [12] Wood, G.M. Slack, P.T. Rossell, J.B. Mann, P.J. P.J. Famell, 1994. J. Agric. Food Chem. 42, 2525.
- [13] Mannina, L., Patumi, M., Fiordiponti, P., Emanuele, M.C., Segre, A.L. 1999c. Olive and hazelnut oils: a study by high-field ¹H NMR and gas chromatography. Ital. J. Food. Sci. 11, 139–149.
- [14] Woodburg, S.E. Evershed, R.P. Rossell, J.B. 1998. J. Am. Oil Chem. Soc. 75, 371.

- [15] Sedrati, S. 2004. Huile d'argane : Composition et intérêt thérapeutique diététique et cosmétique. Fac de Pharmacie, Thèse univ. Mohammed V. 92.
- [16] Ranalli, A. De Mattia, G. Ferrante, M.L. et Giansante, L. 1997. Incidence of olive cultivation area on the analytical characteristics of the oil. Note 1. La Rivista Italiana delle Sostanze Grasse, 74,501-508.
- [17] Aparicio, R. et Luna, G. 2002. Characterization of monovarietal virgin olive oils. European Journal of Lipid Science and Technology, 104, 614-627.
- [18] Aranda, F. Gómez-Alonso, S. Rivera del Alamo, R.M. Salvador, M.D. et Fregapane, G. 2003 Triglyceride, total and 2-position fatty acid composition of Cornicabra virgin olive oil: Comparison with other Spanish cultivars. Food Chemistry, 86, 485-492.
- [19] Zarrouk, M. Marzouk, B. Ben Milled, D. et Cherif, A. 1996. Accumulation de la matière grasse de l'olive et effet du sel sur sa composition. Olivae, 61 ,41-46.
- [20] Abaza, L. Ben Temime, S. M'Sallem, M. Daoud, D. Zarrouk, M. et Cherif, A. 2003. Etude comparative de la lipogénèse chez quelques variétés d'oliviers cultivées en Tunisie. Riv. Ital. Dell Sost. Gr., 80, 297-306.
- [21] Baccouri, B. Ben Temime, S. Taamalli, W. Daoud, D. M'Sallem, M. et Zarrouk, M. 2007. Analytical characteristics of virgin olive oils from two new varieties obtained by controlled crossing on Meski variety. Journal of Food Lipids, 14, 19-34.
- [22] Ollivier, D. 2003. Recherche d'adultération dans les huiles végétales : application à la qualité des huiles vierges et notamment de l'huile d'olive Oléagineux, Corps Gras, Lipides. 10, Numéro 4, 315-20, Problématiques actuelles dans le domaine de l'analyse des oléagineux et des corps gras.
- [23] Maata, N. Kartah, B. Harhar, H. Gharby, S. Benazzouz, E. Guillaume, D. Charrouf, Z. 2011. Détection de l'adultération de l'huile d'argane par des huiles végétales vierges et raffinées Actes du Premier Congrès International de l'Arganier.

Chapitre 7: APPLICATION DE LA SPECTROSCOPIE INFRAROUGE ET LES METHODES CHIMIOMETRIQUES POUR LA DETECTION DE L'ADULTERATION DES HUILES

Partie A : Détection de l'adultération de l'huile d'olive par l'huile de soja et l'huile de tournesol

1. Introduction

La falsification des produits alimentaires implique le remplacement d'ingrédients d'un coût élevé par des substituants moins chers et de qualité inférieure [1]. Il s'agit d'un type de fraude économique qui peut constituer une menace majeure pour la santé des consommateurs. L'huile d'olive extra vierge présente d'excellentes propriétés sensorielles et nutritionnelles qui ont fait l'objet d'une falsification par différents types d'huiles végétales depuis longtemps. L'adultération de l'huile a reçu beaucoup d'attention dans les années 1980 après la consommation de l'huile toxique provoquant plus de 400 morts et 20.000 malades [2-3].

L'importance de trouver un moyen facile, rapide pour détecter la falsification des huiles et de garantir sa qualité est une des préoccupations majeures pour l'industrie agroalimentaire.

Il existe plusieurs méthodes pour la détection de l'adultération comme l'indice d'iode, l'indice de saponification, l'indice de réfraction [4]. En outre, les UV et la spectrométrie de fluorescence ont été également utilisés [4].

L'application des techniques de la spectroscopie IR et Raman combinées avec des méthodes chimiométriques est une approche relativement nouvelle pour déterminer l'authenticité de l'huile d'olive. L'utilisation de méthodes spectroscopiques pour la détection de l'authentification d'huiles végétales a apparu dès les années 90 [5-7].

La chimiométrie, un sujet traitant de l'analyse multivariée des données, a contribué et aidé à résoudre beaucoup de problèmes dans l'analyse qualitative et quantitative.

Les techniques spectroscopiques, MIR- infrarouge Raman et proche infrarouge PIR ont la possibilité de quantifier l'adultération de grignons d'olive dans de l'huile d'olive extra - vierge (0 % et 100 % en incréments de 5% en poids) avec la valeur R^2 supérieur de 0,99 [8] pour chaque technique. Christy et al . [9] ont étudié la spectroscopie NIR pour détecter et quantifier l'altération de l'huile d'olive avec les huiles du soja , de tournesol , de maïs , de

noix et de noisettes qui ont abouti à des limites d'erreur de $\pm 0,5\%$ (huile de maïs), $\pm 1,32\%$ (tournesol), $\pm 0,96\%$ (huile de soja) $\pm 0,56\%$ (huile de noix) et $\pm 0,57\%$ (huile de noisette). La spectroscopie IR à Transformée de Fourier (FTIR) équipée d'un accessoire ZnSe - ATR était capable de détecter l'adultération de l'huile d'olive vierge avec de l'huile de noisette au taux de 25 % et plus [10].

L'objectif principal de cette étude était de construire un modèle multivarié en mesure d'expliquer et de prédire les changements observés dans la concentration du soja et adultérants de tournesol (variables y_i) en fonction des spectres MIR mesurée (bloc X de variables). Il s'agit de trouver une relation mathématique entre ces deux ensembles de données de variables. En particulier, dans ce travail, l'Analyse en Composantes Principales (ACP) et la régression au sens des moindres carrés Partiels (PLSR) ont été évalués dans la modélisation de la concentration des adultérants.

2. Matériel et méthode

2.1. Préparation des échantillons

Les huiles végétales de soja et de tournesol ont été achetées aux prés des supermarchés locaux. Les compositions du mélange binaire huile d'olive - huile de soja (58 échantillons) et huile d'olive -huile de tournesol (46 échantillons) a été déterminée à partir du pourcentage massique, le domaine de la falsification est de 0 à 24% en masse.

Cette méthode est considérée comme méthode de référence. En effet, la qualité du modèle de prédiction dépend non seulement de la qualité du spectrophotomètre mais aussi de celle de la méthode de référence [11].

2.2. La Spectroscopie Infrarouge

Le spectromètre IR à Transformée de Fourier utilisé pour cette étude est le Bruker (Vector 22) permettant l'acquisition de spectres dans le moyen infrarouge en fonction de la combinaison source/séparatrice/détecteur utilisée.

Le Vector 22 est piloté par un logiciel d'acquisition et de traitement de données "OPUS version 5.0.

Paramètres de la méthode

❖ Paramètres optiques

Notre spectromètre IR est équipé d'une source de Globar (MIR) avec une séparatrice en KBr Germanium et un détecteur de type DTGS.

❖ Paramètres d'acquisition

Spectre résultant: Absorbance, résolution est de 4 cm^{-1} , accumulations pour l'échantillon (nombre de "scans") est de 124, accumulations pour la référence est de 124, limite supérieure d'acquisition est de 4000 cm^{-1} et la limite inférieure d'acquisition est de 400 cm^{-1} .

La base de données utilisée pour le développement des différents modèles de prédiction de la composition chimique, à partir des spectres moyen infrarouge, est composée de 104 mélanges.

Les 104 spectres Moyen Infrarouge MIR obtenues ont été converties en données Matlab et ont été utilisés pour l'accumulation du bloc de variables X (matrice de données X). Les concentrations des différentes adultérations (huiles de soja et de tournesol) ont été utilisées comme bloc de variables y. Deux lots de données ont été préparés, l'un pour construire le modèle de calibration (92 mélanges) et un autre lot de validation externe (12 mélanges) qui seront consacrés à un test indépendant.

La PCA est une méthode multivariée basée sur l'exploration de données spectrales. Les premières travaux est préliminairement effectués par Jolliffe 1986 [12].

Il s'agit d'un procédé de projection, ce qui donne des informations sur la structure de l'ensemble de données latentes. Il transforme un grand nombre de variables (peut-être) d'origine corrélées dans un plus petit nombre de variables non corrélées, les variables orthogonales expliquant la variance maximale appelée composantes principales. Des lots de charges variables en composantes principales permettent l'interprétation des principales sources de variance des données dans les spectres. Le principal avantage de l'ACP est de réduire la dimensionnalité du problème (nombre de variables), mais conserve la majeure partie de la variabilité d'origine dans les données expérimentales et des filtres de bruit et sources mineures de variance. Par conséquent, PCA permet une interprétation plus simple de la variance des données spectrales.

La régression PLSR permet également de maximiser la covariance entre X et y. La PLSR recherche pour le sous-espace de facteur le plus congruent à deux matrices, et ses prédictions sont généralement mieux que d'utiliser d'autres méthodes de régressions multilinéaires. Une nouvelle matrice de pondération (qui reflète la structure de covariance entre les prédicteurs X et les variables de réponse y) est calculée et fourni une riche information pour l'interprétation des facteurs.

La sélection de nombre optimal de composantes (variables latentes) dans PLSR a été effectuée en utilisant une validation croisée interne ((leaving-one-out) et de prédiction optimale des valeurs de y. Récemment, un nouvel outil puissant pour l'interprétation de la complexité des modèles de PLSR (construits avec plusieurs variables latentes), a été démontré [13].

Au cours de cette étude, il a été exploré différentes régions spectrales de la gamme spectrale complète 4000-600 cm^{-1} constituant la matrice X. Il a été trouvé sur la base de la meilleure capacité de prédiction obtenue que la zone spectrale la plus importante est 1495-675 cm^{-1} .

3. Résultats et discussion

3.1. Attribution des bandes d'absorption de l'huile d'olive dans le moyen infrarouge

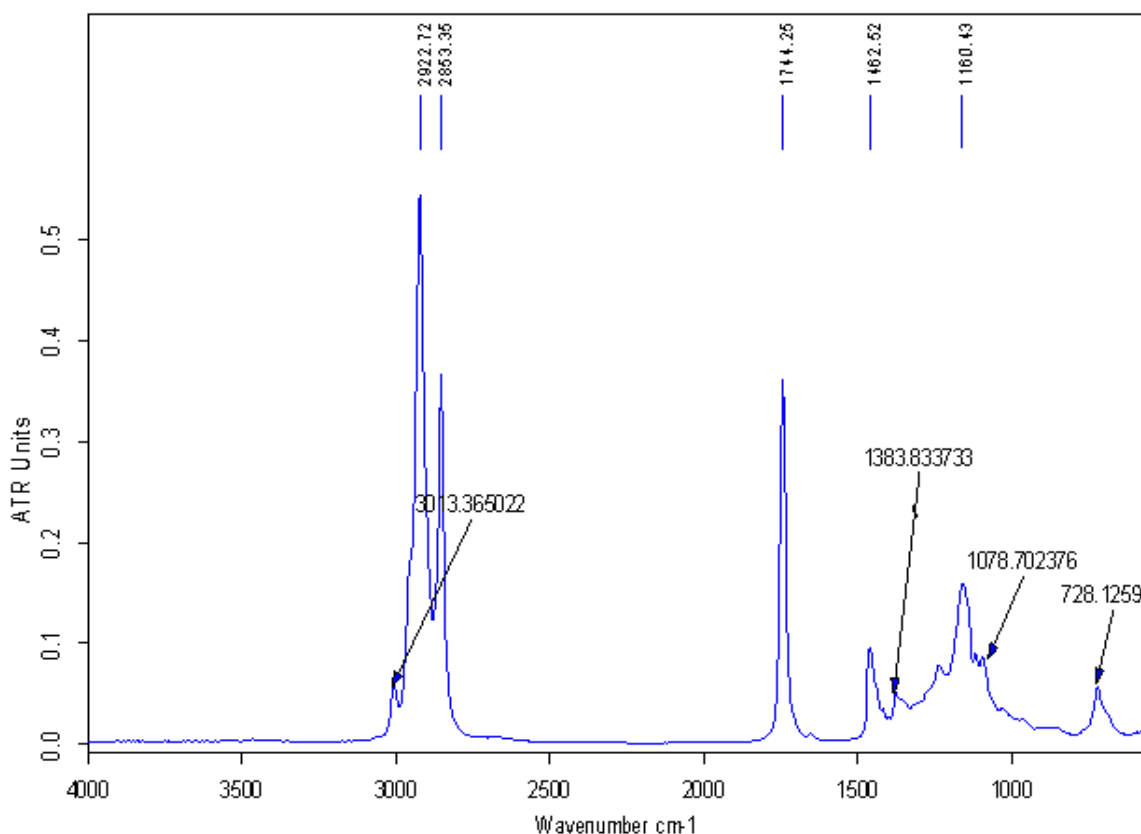


Figure 21 : Spectre moyen infrarouge de l'huile d'olive

Dans le spectre de la figure 21 nous observons les bandes suivantes :

- Une bande à 3013 cm^{-1} attribuée à l'élongation des groupements =C-H.
- La zone 3000-2800 cm^{-1} est composée par :

-Une bande à 2980 cm^{-1} d'intensité faible est associée à l'élongation symétrique du groupe CH_3 .

-Une bande à 2961 cm^{-1} attribuable à la vibration d'élongation asymétrique du groupement CH_2 .

-Une bande intense à 2922 cm^{-1} attribuable à la vibration d'élongation asymétrique du groupement CH_3 .

-Une bande à 2854 cm^{-1} est associée à l'élongation symétrique du groupement CH_2 .

➤ $1700\text{-}1755\text{ cm}^{-1}$:

-L'élongation du groupe $\text{C}=\text{O}$ des esters est visible vers 1743 cm^{-1} .

➤ $1500\text{-}1300\text{ cm}^{-1}$:

-Les bandes à 1462 cm^{-1} et 1383 cm^{-1} sont attribuables respectivement aux vibrations de déformation des groupes CH_2 et CH_3 .

➤ 1160 cm^{-1} cette bande est attribuable aux vibrations d'élongation du groupe $\text{C}-\text{O}$ ester.

➤ Une bande à 1050 cm^{-1} est attribuée à la bande $\text{O}-\text{C}-\text{C}$ des esters dérivés des alcools primaires.

➤ Une bande à 728 cm^{-1} , l'intensité de cette bande est directement liée aux vibrations de déformation du groupe CH_2 .

3.2. Huile d'olive, huile de soja et huile de tournesol

La figure 22 montre les spectres moyens infrarouges pour les trois huiles, l'huile d'olive (bleu), huile de soja (rose) et l'huile de tournesol (rouge) pures.

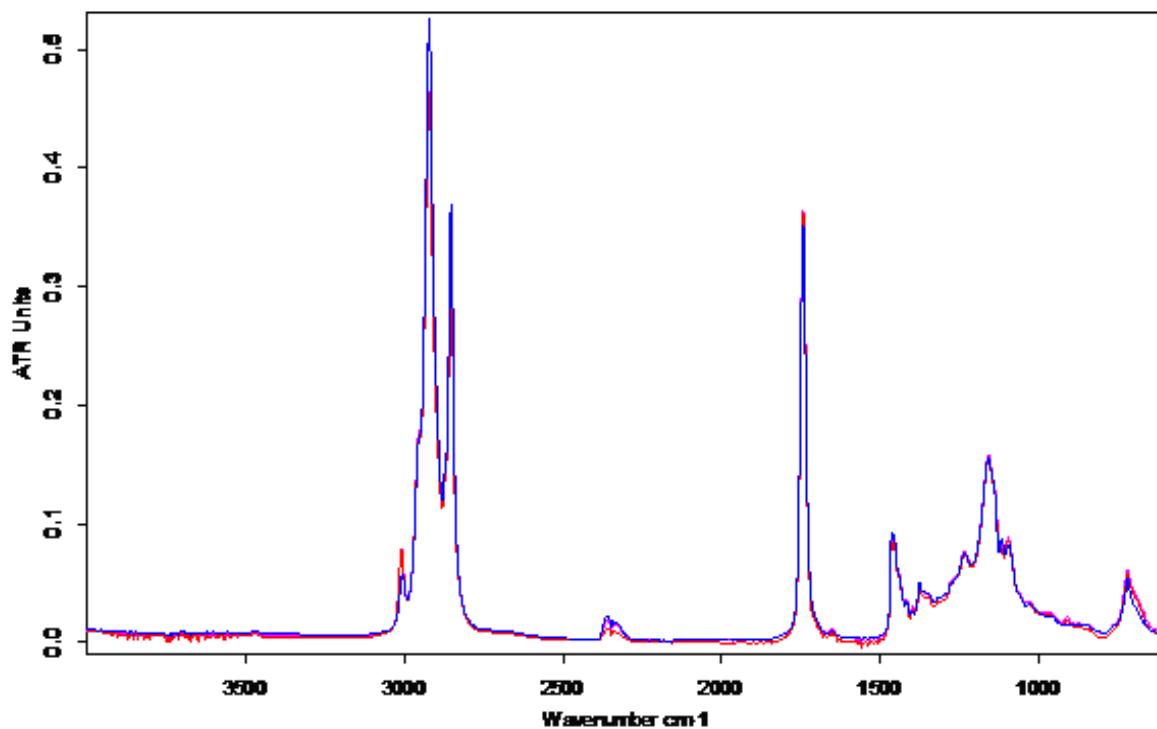


Figure 22 : Spectres MIR de l'huile d'olive (bleu), huile de soja (rose) et l'huile de tournesol (rouge)

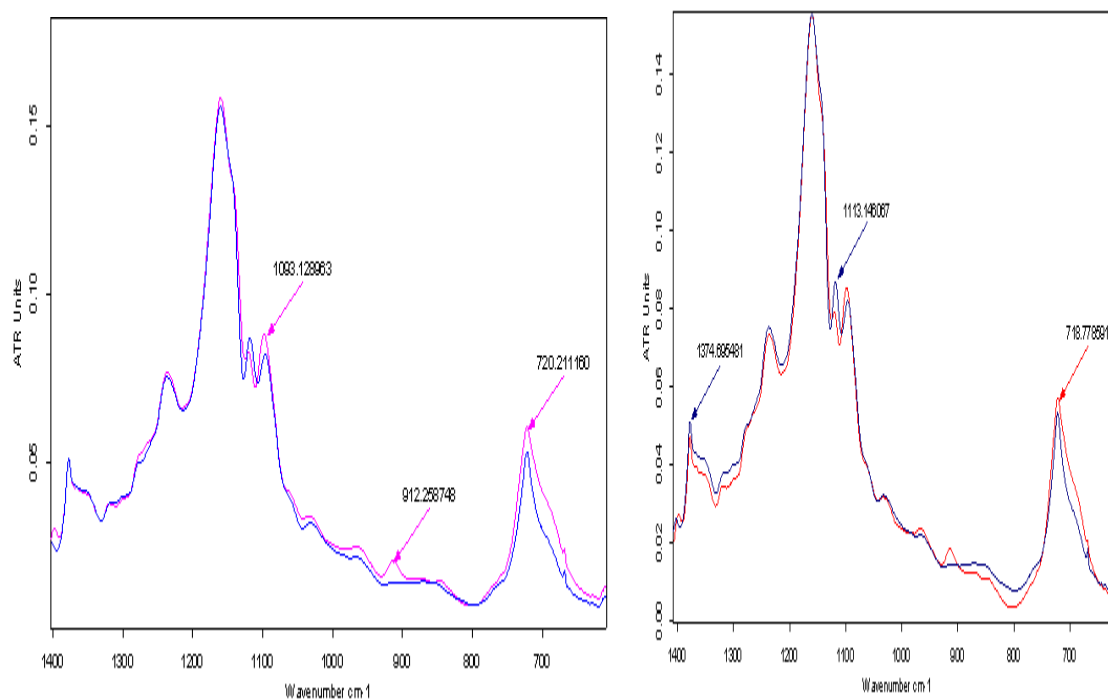


Figure 23 : les trois spectres moyens infrarouges de l'huile d'olive de l'huile de soja et l'huile de tournesol purs.

La figure 23 montre les trois spectres moyens infrarouges de l'huile d'olive de l'huile de soja et l'huile de tournesol purs.

Les trois spectres de l'huile d'olive, l'huile de soja et l'huile de tournesol se ressemblent, la seule différence existant entre les trois spectres réside dans l'intensité des pics dans la zone spectrale de (1400-700 cm^{-1}). Cela est dû à la composition chimique similaire des huiles.

L'information spectrale utilisée pour l'analyse discriminante est contenue dans les régions de nombre d'onde de 3100-2800 cm^{-1} et 1800-900 cm^{-1} .

Analyse chimiométrique

Récemment, un nouvel outil puissant pour l'interprétation de la complexité des modèles de PLSR (construits avec plusieurs variables latentes), a été démontré [13]. L'interprétation de l'importance variable de la projection, les scores de VIP, de modèle de régression est utile pour l'évaluation de l'importance de chaque variable dans la projection utilisée dans un modèle de PLSR et est souvent utilisé pour la sélection de variables. Une variable avec un score de VIP près ou supérieure à un (1) est considérée comme importante dans un modèle donné.

3.2 Analyse en composante Principale

L'analyse en composantes principales (ACP) recherche les directions de variances maximales, à partir de la matrice de corrélations. La décomposition en valeurs propre de cette matrice crée une nouvelle base vectorielle. PCA a été appliquée pour différencier entre les deux mélanges.

Tableau 14: le % de la variance pour le modèle ACP

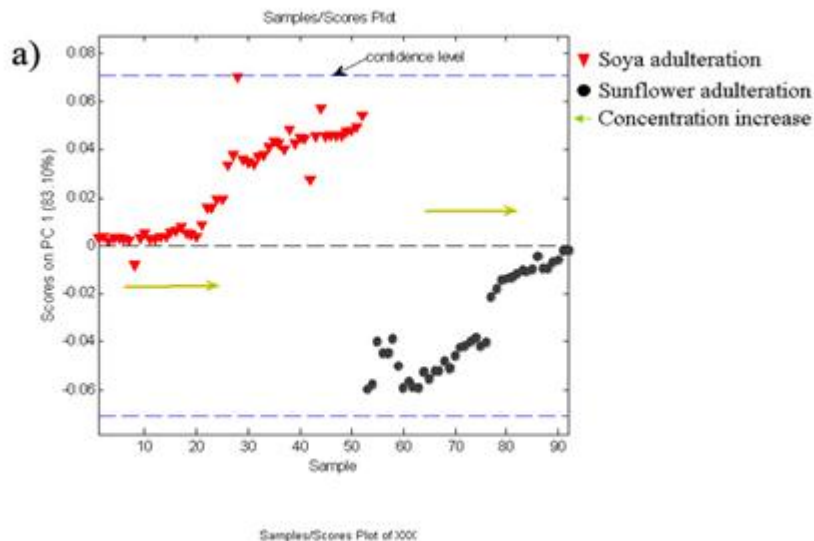
| Principal component Number | Eigen value of Cov(X) | % Variance Captured this PC | % Variance Captured Total |
|----------------------------|------------------------|-----------------------------|---------------------------|
| 1 | 1.27 e ⁻⁰⁰³ | 83.10 | 83.10 |
| 2 | 1.27 e ⁻⁰⁰⁴ | 8.27 | 91.37 |
| 3 | 7.83 e ⁻⁰⁰⁵ | 5.11 | 96.48 |

Le modèle a été en mesure de classer jusqu'à 23,84 en pourcentage massique l'adultération par l'huile de tournesol et de l'huile de soja. La dimension du modèle a été réduite en choisissant des valeurs propres (CP).

Le tableau 15 montre que 96,48 % de l'information spectrale a été expliquée par les trois composantes principales. La 1^{ère} composante explique 83.10% de la variance totale on dit que la 1^{ère} CP explique au mieux les variables, tandis que la 2^{ème} composante explique 8.27% des mélanges binaires.

3.3. Partial least square Régression

Dans le but de quantifier la concentration de l'huile adultérée dans des échantillons d'huile d'olive, la technique de régression PLS a été utilisée. Le domaine spectral utilisé dans l'analyse PCA est la même pour la régression PLS. La gamme de concentration étudiée était 0-24% de la falsification. Les données ont été centrées pour réduire le bruit instrumental et la variation systématique avant la régression PLS. La sélection du nombre optimal de composantes (variables latentes) dans le PLSR a été effectuée en utilisant une validation croisée interne « leave-one-out ». On choisit le nombre de composantes latentes de manière à avoir la plus petite erreur résiduelle pour la prédiction optimale des valeurs de y.



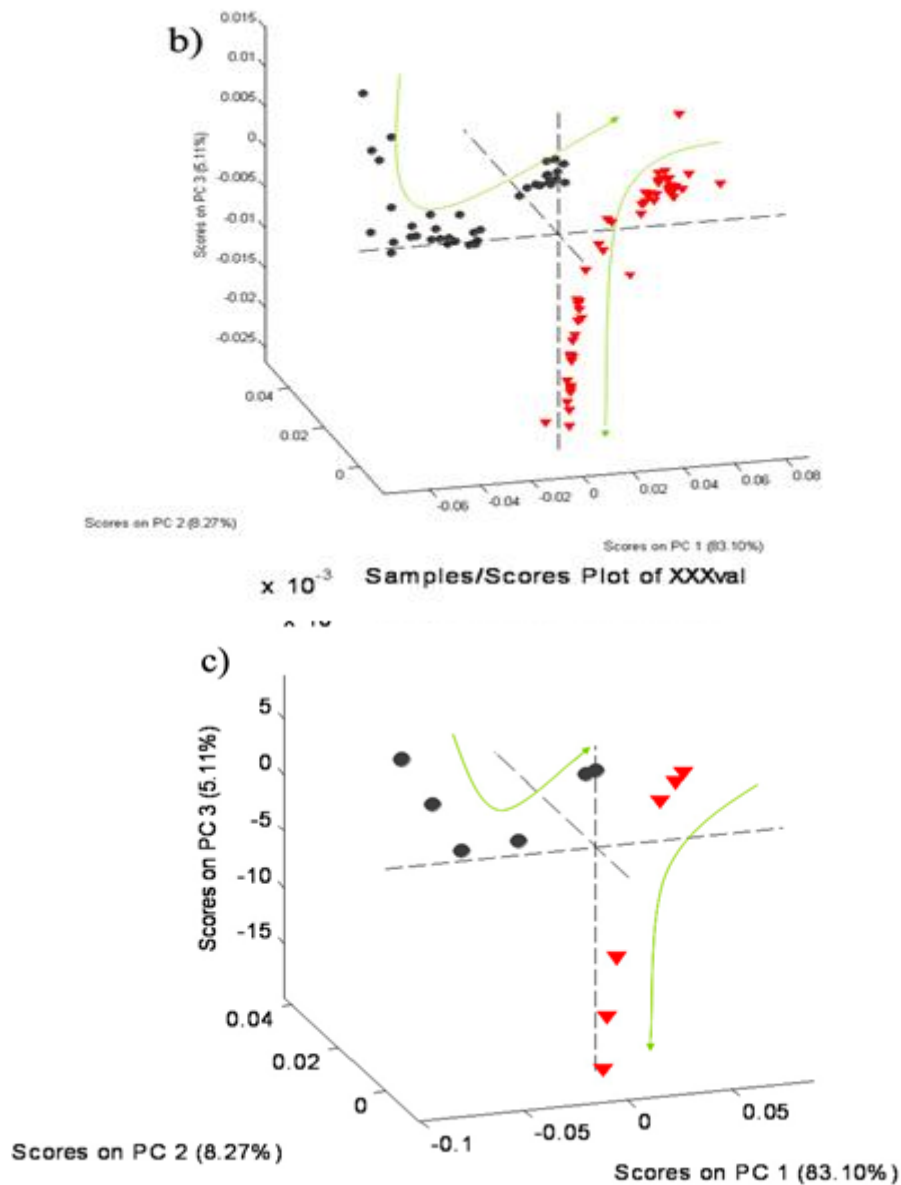


Figure 24 a) score plot de la 1^{ère} principale composante pour la classification des mélanges binaires huile de soja –huile d’olive et de huile de tournesol-huile d’olive ; b): 3D scores plot des trois premiers PC de la classification des mélanges binaires huile de soja –huile d’olive et de huile de tournesol-huile d’olive à l’étape de l’étalonnage; c) les résultats de la validation externe (Cross validation: leave one out).

La figure 24a montre une inverse corrélation entre les échantillons adultérés par l’huile de soja et ceux adultérés par l’huile de tournesol. La 1^{ère} composante principale est responsable de la classification des deux mélanges binaires. La 1^{ère} composante n’explique pas bien la quantité d’adultérations des huiles, par exemple une concentration faible du soja a la même influence sur la 1^{ère} composante qu’une concentration élevée en tournesol.

D'autre part, l'augmentation de l'adultération par l'huile de soja augmente avec l'augmentation de cet échantillon à un mélange contenant une concentration faible de l'huile de tournesol, pour ce fait nous proposons qu'un graphique en 3D puisse être plus approprié pour expliquer la différence de falsification par rapport au profil de concentration de la substance étrangère. La figure 24b montre la dispersion des échantillons en fonction de leur type et de la concentration de l'adultérant où les trois dimensions représentent les 1^{ère}, 2^{ème} et 3^{ème} vecteurs propres.

La figure 24b montre l'importance relative de chaque adultérant (soja ou huile de tournesol) avec une concentration différente ; ce qui contribue à la différence de la variance de chaque composante principale dans l'espace tridimensionnel 3D. La représentation 3D des coefficients des trois premiers vecteurs propres, permet de séparer les deux populations

En outre le modèle ACP obtenu est appliqué pour valider un lot de 12 mélanges, six échantillons pour chaque binaire.

La figure 24 c montre la capacité du modèle pour la prédiction des douze nouveaux échantillons et de la classification des deux groupes.

Le modèle ayant la plus faible erreur de prédiction a été choisi avec 10 variables latentes expliquant plus de 99% de l'information spectrale donnant également la plus faible erreur résiduelle de validation croisée interne 0,37 (Figure 23a)

La figure 25a montre les valeurs de concentration prédites obtenues à partir du modèle PLS en fonction des concentrations réelles des huiles de soja et de tournesol dans les échantillons d'huile d'olive au cours de l'étape de calibration. La valeur élevée de R^2 révèle une très forte corrélation entre les valeurs mesurées et les valeurs prédites de l'adultérant dans l'huile d'olive.

Le modèle a un nombre élevé de variables latentes 10 LV, vu que le point critique est le choix du nombre optimal de variables latentes (LVs). La raison pour laquelle le choix du nombre de variables latentes est primordial est le problème d'un "*underfitting*" (lorsque trop peu de LVs sont choisies, elles ne contiennent pas toute l'information des données de départ nécessaires pour correctement prédire y ou d'un "*over-fitting*" (lorsque trop de LVs sont sélectionnées, ils existent parmi elles des LVs non-informatives mais contenant du bruit qui risque d'être pris en compte dans le calcul des valeurs à prédire). Par conséquent, on a recours toujours à des méthodes de validation telles que la *Validation Croisée* et le *critère de Durbin-*

Watson pour déterminer le nombre optimal de LVs. Pour ce faire l'étape de validation externe était nécessaire.

3.4. Validation du modèle

Le modèle PLS avec 10LVs a été choisi pour le modèle de validation externe pour être appliquée sur des nouveaux mélanges (au nombre de 12) Figure 25b.

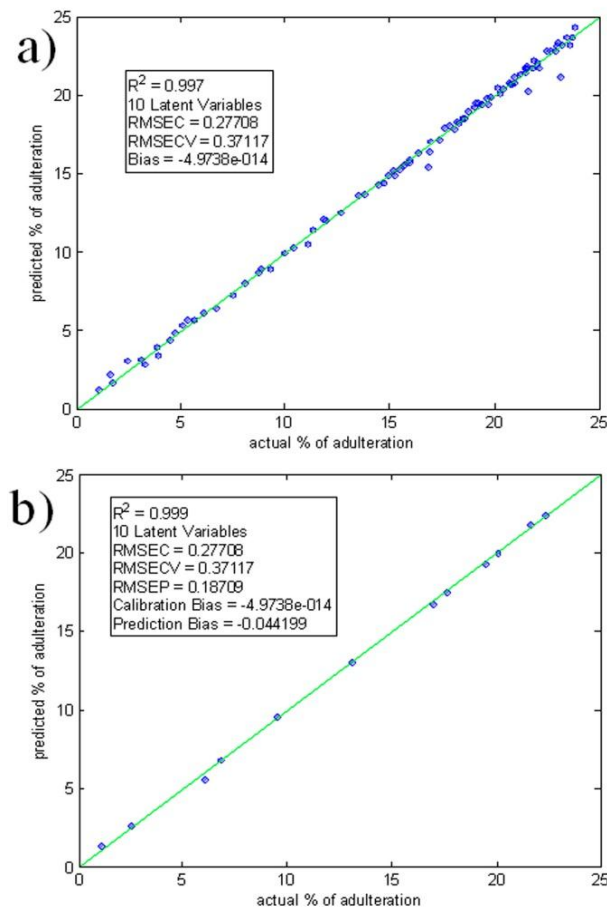


Figure 25: les valeurs prédites par le modèle PLS en fonction des valeurs mesurées a) pour le modèle de calibration b) pour le modèle de validation externe

La figure 25a montre les rapports entre les valeurs prédites et les valeurs mesurées de l'huile de soja et l'huile de tournesol dans l'huile d'olive pour les échantillons d'étalonnage. Cette figure montre que une très forte corrélation entre les valeurs mesurées et les valeurs prédites de l'adultérant dans l'huile d'olive a été obtenue.

La figure 25b montre une bonne corrélation entre les valeurs prédites et les valeurs mesurées avec R^2 de 99,9 %, RMSEP égale à 0,18. Ce résultat confirme la robustesse de la spectroscopie moyen infrarouge pour prédire la qualité de l'huile d'olive extra-vierge, en terme de taux d'adultération.

Le bon chiffre obtenu dans la calibration et la validation du modèle de PLS dans la région spectrale analysée décrit la bonne capacité prédictive de ce modèle et a démontré sa qualité. En outre, il a souligné la pertinence des applications de la méthode proposée pour la détermination de la falsification similaire.

3.5. L'importance variable en projection ou VIP score

En outre, l'étude a porté sur les variables responsables des résultats obtenus, par exemple la partie de la gamme spectrale utilisée. Un nouvel outil puissant pour l'interprétation de la complexité des modèles de PLSR (construits avec plusieurs variables latentes), a été appliqué dans cette étude.

L'interprétation de l'importance variable en projection ou VIP scores obtenue à partir du modèle PLS a fait apparaître l'importance de chaque variable (nombre d'onde en cm^{-1}).

L'étude de l'évaluation variable a été suivie de la règle générale pour la sélection de variables. Si une variable avec un score de VIP est proche ou supérieure à un (1), elle est considérée comme importante dans un modèle donné.

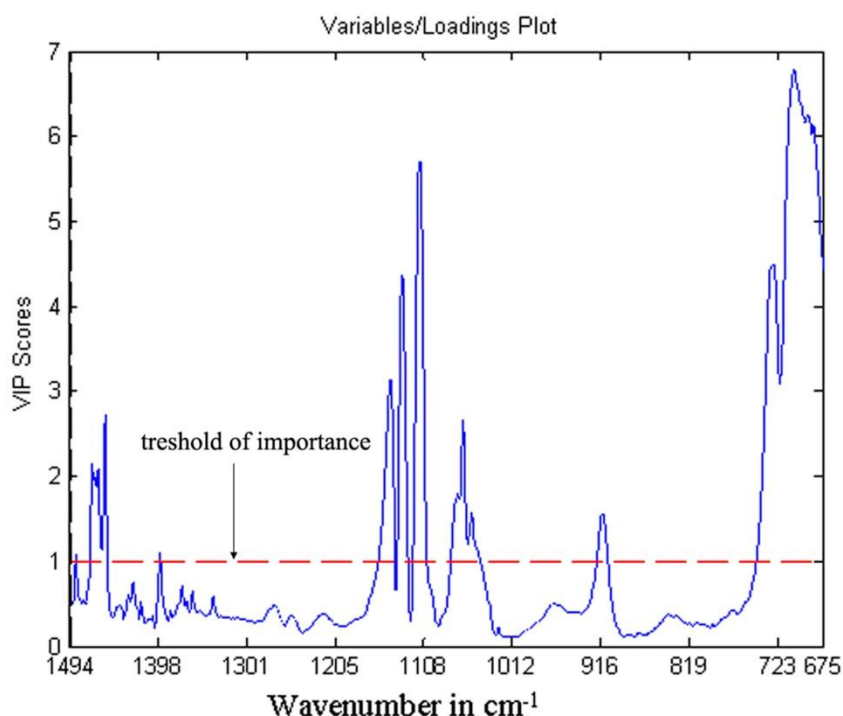


Figure 26 : L'importance variable en projection ou VIP scores pour la région spectrale étudiée $1494\text{-}675 \text{ cm}^{-1}$ lorsque des mélanges binaires huile d'olive –huile de soja et huile de tournesol- huile d'olive ont été modélisés simultanément par la PLS.

La figure 26 montre les scores de projection obtenus pour la région spectrale étudiée 1494-675 cm^{-1} lorsque des mélanges binaires : huile de soja -huile d'olive et huiles d'olive - huile de tournesol ont été modélisés simultanément par PLS. Trois zones spectrales majeures peuvent être distinguées comme très importantes pour la performance du modèle PLS. La première est entre 1473,3 et 1455 cm^{-1} ; la deuxième se situe entre 1160 à 1106 cm^{-1} avec quelques exceptions et la troisième zone a été constatée entre 746,32 et 674,96 cm^{-1} , où le domaine spectral étudié dans cette étude a été terminé.

Le tableau 15 résume les régions spectrales où les VIP ont été jugées sur le seuil de un.

Tableau 15 : les régions spectrales où les VIP ont été jugées sur le seuil de un.

| | |
|--------|--------|
| 1489.7 | 1488.8 |
| 1473.3 | 1455 |
| 1398.1 | 1396.2 |
| 1160 | 1141.7 |
| 1137.8 | 1126.2 |
| 1120.4 | 1106 |
| 1079.9 | 1048.1 |
| 920.84 | 909.27 |
| 746.32 | 674.96 |

Conclusion

La détermination de l'authenticité de l'huile d'olive extra vierge est devenue plus importante ces dernières années. L'étude a porté sur la détection et la quantification de la falsification de l'huile d'olive extra-vierge avec l'huile de soja et l'huile de tournesol en utilisant l'infrarouge moyen (MIR) traité par des outils chimiométriques.

La présence de l'huile de tournesol et l'huile de soja (comme adultérant) dans l'huile d'olive extra vierge est détectée par la spectroscopie FTIR en utilisant l'ATR.

Cette étude démontre l'application de l'analyse en composantes principales PCA afin de faire la distinction entre les huiles d'olive « falsifiées par des huiles de soja et de tournesol en utilisant la spectroscopie infrarouge et les méthodes chimiométriques lorsque deux données de mélanges binaires sont dans une matrice de données. Les deux premiers PC permettent la

classification des deux mélanges et utilisent suffisamment d'informations pour expliquer le modèle, mais certaines complications liées à la concentration ont été surmontées quand un graphique en 3D des scores a été utilisé à la fois dans l'étalonnage et l'étape de validation externe.

En outre, une nouvelle approche PLS a montré qu'elle est un outil très puissant permettant la compréhension du processus de calibration du modèle pour l'analyse quantitative des mélanges binaires. Une très bonne corrélation a été obtenue avec un coefficient R^2 de 0.997, un faible RMSEV 0.37117 dans la calibration et une erreur moyenne pour la prédiction de l'ordre de -0,0506.

Le succès de l'application du modèle final montre la capacité de la chimiométrie comme technique appliquée sur les données MIR-FT pour une rapide et précise classification des huiles d'olive sans la nécessité d'une préparation préalable des échantillons.

Références bibliographiques

- [1] Tay, A., Singh, R.K. Krishnan, S.S. Gore. J.P 2002. Authentication of olive oil adulterated with vegetables oils using Fourier transform infrared spectroscopy. LWT-Food Science and Technology. 35, 99-103.
- [2] Jimeno S.A., 1982. The Spanish toxic symptoms. Trends anal Chem. 1, 4-6
- [3] Posada M, Castro M, Kilbourne EM, Diaz de Rojas F, Abaitua I, Tabuenca JM, Vioque A. 1987. Toxic –oil syndrom:case. Reports associated with the ITH oil refinery in Sevilla. FOOD Chem.Toxicol. 25, 87-90.
- [4] Tous G., 1986 in” Analysis and characterisation of oils, fats and fat products”, ed .H.A.Boekenoogen,.vol.2,interscience,London,315.
- [5] Baeten V., Meurens M., Morales M.T.& Aparicio. R., 1996. Detection of virgin olive oil adulteration by Fourier transform Raman spectroscopy.J.Agricultural Food Chemistry ,44 (8), 2225-2230.
- [6] Law Y.W., Kemsley E.K.,& Wilson R.H. 1994. Potential of fourier transform infrared spectroscopy for the authentication of vegetable oils. J.Agricultural Food Chemistry, 42, 1154-1159.
- [7] Law Y.W., Kemsley E.K. & Wilson R.H. 1995. Potential of fourier transform infrared spectroscopy for the authentication of vegetable oils. J.Agricultural Food Chemistry, 53, 95-98.
- [8] Yang H., & Irudayaraj J., 2001. Comparison of near-infrared, Fourier transform- infrared, and Fourier transform-Raman methods for determining olive pomace oil adulteration in extra virgin olive oil. J.of the American Oil Chemists society, 78, 889-895.
- [9] Christy A.A., Kasemsumran S., Du Y.& Ozaki Y., 2004. The detection and quantification of adulteration in olive oil by near-infrared spectroscopy and chemometrics, J. Analytical Sciences, 20, 935-940.
- [10] Ozen B.F., & Mauer L.J, 2002. Detection of hazelnut oil adulteration using FT-IR spectroscopy. J. Agricultural Food Chemistry, 50, 3898-3901.

[11] Bernard D. 2000. La spectroscopie infrarouge et ses applications analytiques, TEC&DOC, Londres-Paris-New York, 3-26.

[12] Jolliffe, T., O.B. Allen, B.R. Christie, 1989. Comparison of variety means using cluster analysis and dendrograms, *Experimental Agriculture*, 25, 259-269.

[13] Il-Gyo Chong and Chi-Hyuck Jun. 2005. Performance of some variable selection methods when multicollinearity is present *Chem. Intell. Lab. Sys.*, 78,103-112.

Partie B : Détection de l'adultération de l'huile d'argan par l'huile de soja et l'huile de tournesol

1. Introduction

La grande différence entre les prix de l'huile d'argan et les autres huiles végétales vierges et raffinées peut entraîner des tentations de falsification.

Une falsification peut consister à introduire des huiles de moindre coût dans l'huile d'argan. A cet effet, l'étude de mélanges d'huile d'argan avec d'autres huiles végétales vierges telles que les huiles de soja et de tournesol a été effectuée.

L'authentification des échantillons d'huile d'argan nécessite généralement l'utilisation de techniques d'analyse sophistiquées et ne demandent pas beaucoup de temps. La détermination de l'authenticité de l'alimentation et la détection de l'altération sont des problèmes de plus en plus importants dans l'industrie alimentaire. Il est nécessaire pour les tests de contrôle de qualité des techniques d'analyse simples et rapides. La détermination de la falsification d'argan par des huiles végétales est un défi analytique réel. Cette étude se concentre sur la détection et la quantification de l'huile d'argan adultérée par différentes huiles comestibles en utilisant la spectroscopie moyen-infrarouge (MIR) et la chimométrie.

La variation chimique entre l'huile extra vierge et d'autres huiles nécessite une très haute qualité spectrale imposant des exigences expérimentales rigoureuses, afin d'appliquer avec succès l'analyse discriminante. Cependant, Lai et al [1] ont été en mesure de développer avec succès un modèle discriminant et de fournir un procédé approprié pour une utilisation dans un temps d'analyse d'environ 2-3 min.

Les techniques de spectroscopie, telles que la spectroscopie moyen infrarouge (MIR) ou le proche infrarouge (NIR) ont montré leur utilité pour le contrôle de la qualité des huiles d'olive [2-3], par exemple l'évaluation rapide des altérations de l'huile d'olive vierge [4-5]. Pour l'analyse qualitative, le moyen infrarouge (MIR) a plus d'applications parce que les empreintes digitales des groupes fonctionnels sont situées dans la région du MIR (4000 - 400 cm^{-1}).

La spectroscopie infrarouge à transformée de Fourier (FTIR) avec totale réflectance atténuée (ATR) ou Accessoires de cellulaires de transmission a été utilisée pour authentifier, identifier ou classer les graisses et les huiles [6-8]. Jusqu'à présent, peu d'études sont disponibles sur l'adultération de l'huile d'argan par des huiles comestibles [9] et la majorité de ces études ont été effectuées par chromatographie en phase gazeuse. L'objectif principal de

cette étude est de démontrer le potentiel de la spectroscopie infrarouge pour la détermination quantitative des niveaux d'adultération de l'huile d'argan par l'huile de tournesol et l'huile de soja.

La méthode utilise la spectroscopie FTIR, la régression en moindres carrés partiels (PLS) et l'analyse en composantes principales (ACP).

2. Matériel et méthodes

Les huiles d'argan, de tournesol et de soja ont été obtenus à partir des marchés locaux. L'huile d'argan a été adultérée avec de l'huile de tournesol et l'huile de soja à des niveaux d'adultération compris entre 0 et 30% en pourcentage massique (p / p). Un total de 67 mélanges binaires de tournesol –huile d'argan et 37 mélanges binaires argan- soja ont été préparés avec 53-29 attribués pour une utilisation en tant que données de calibration et des lots de 14-8 pour la validation pour les huiles de tournesol et de soja, respectivement.

La Spectroscopie Infrarouge

Le spectromètre IR à Transformée de Fourier utilisée pour cette étude est le Bruker (Vector 22) permettant l'acquisition de spectres dans le moyen infrarouge en fonction de la combinaison source/séparatrice/détecteur utilisée (Les paramètres de la méthode sont les mêmes que celles utilisées dans l'étape 1)

La base de données utilisée pour le développement des différents modèles de prédiction de la composition chimique, à partir des spectres moyen infrarouge, est composée de 104 mélanges.

Les analyses statistiques

Toutes les analyses chimiométriques sont réalisées à l'aide de logiciel UNSCAMBLER (CAMO, Norway) Version 9.6.

Une étape de prétraitements mathématiques des spectres est donc indispensable pour améliorer la qualité du signal et réduire les effets indésirables, différents prétraitement ont été testés. Le nombre optimum de composantes principales est sélectionné en utilisant la validation croisée

3. Résultats et discussion

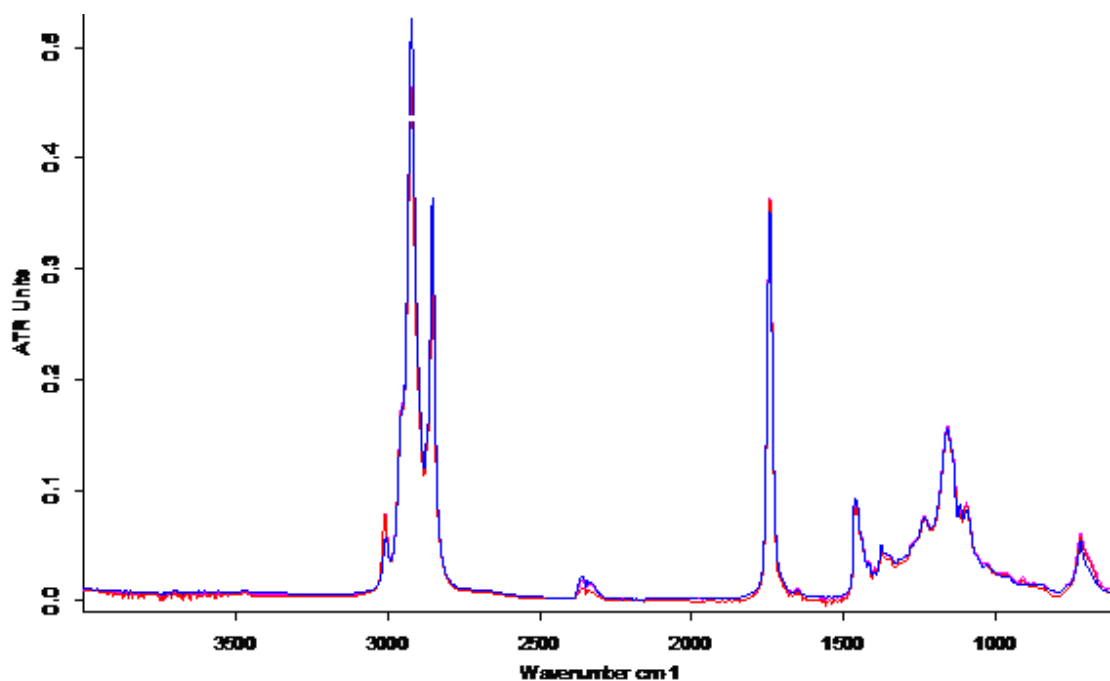


Figure 27 : spectres infrarouges FTIR pour l'huile d'argan (bleu), huile de soja (rouge) et l'huile de tournesol (rose)

La figure 27 montre les spectres infrarouges de l'huile d'argan, huile de soja et l'huile de tournesol

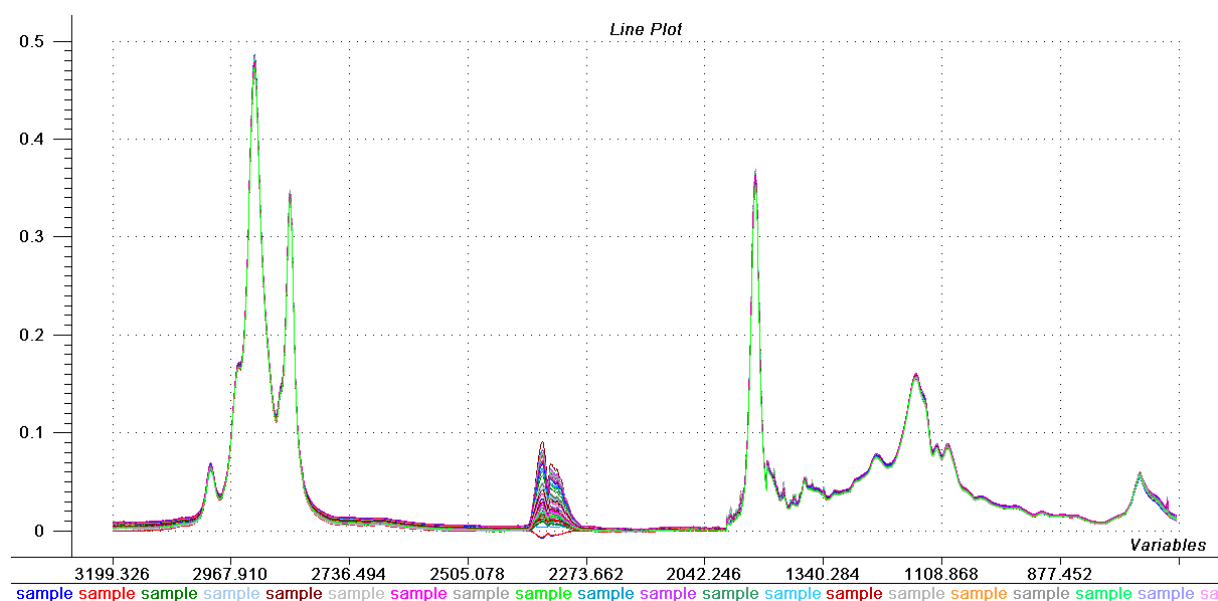


Figure 28 : Les spectres moyens infrarouges des échantillons de l'huile d'argan – huile de tournesol dans la région 3200-600 cm^{-1}

La figure 28 montre les spectres moyen infrarouge de l'huile d'argan adultérée par l'huile de tournesol dans la zone 3200-600 cm^{-1}

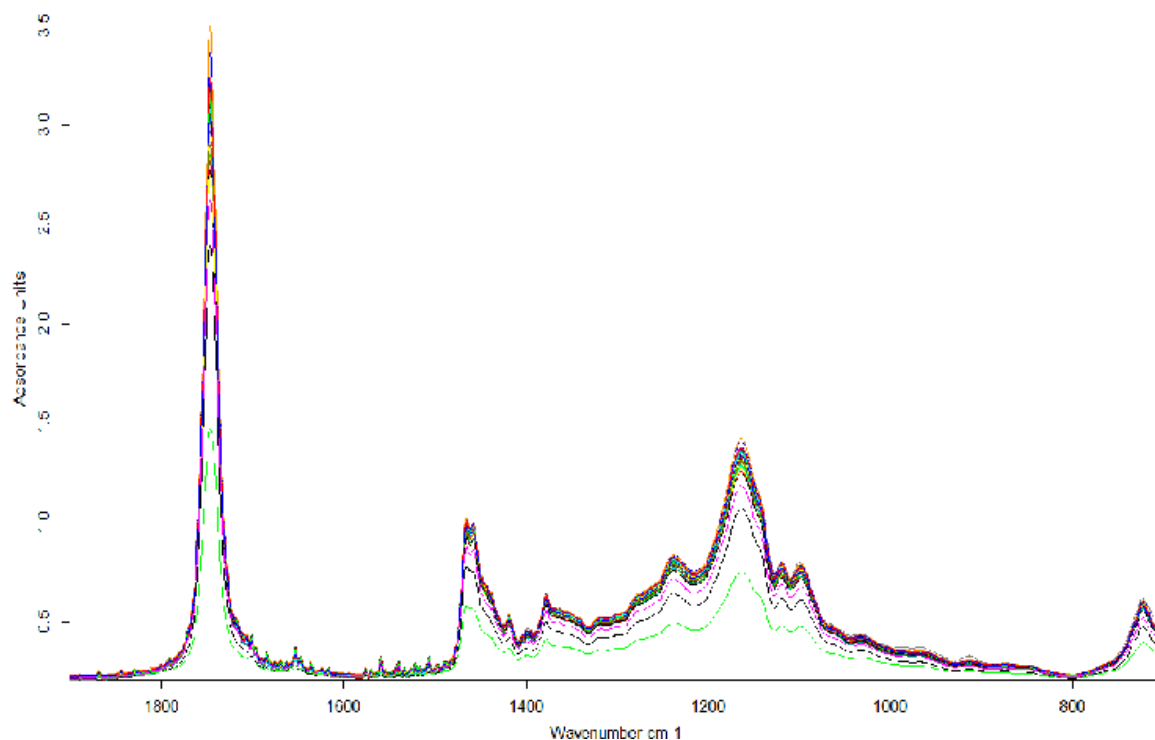


Figure 29 : les spectres moyens infrarouges de l'huile d'argan avec l'huile de soja dans la zone 700-1900 cm^{-1} .

On peut noter qu'aucun pic caractéristique de l'une des deux huiles est n'observé dans cette région du spectre. En outre pour l'analyse de PLS, la région comprise entre 2400 et 2260 cm^{-1} a été supprimée en raison de son faible rapport signal /bruit et de la présence de fluctuation qui est indépendante de la composition de l'échantillon [10-11]. La région entre 4000 et 3030 cm^{-1} a également été supprimé, car il ne contenait pas d'information chimique utile et contribu au bruit instrumental.

La quantification de l'adultération de l'huile d'argan par l'huile de tournesol et l'huile de soja est réalisée en utilisant la régression PLS. Différents prétraitements sur différents zones spectrales ont été testés. La Dérivée est comme le lissage, cette transformation est pertinente. La dérivée de Savitzky –Golay a été choisie pour ces études [12].

Les effets des différents prétraitements sur les spectres des huiles sont visibles généralement sur la dérivée de la ligne de base et sur la dispersion des spectres. Certains prétraitements ont plus d'effet sur les spectres que d'autres.

Une validation croisée est réalisée afin de déterminer le nombre de variables latentes (LVs) à conserver pour construire les modèles.

Pour évaluer les résultats, l'Erreur quadratique moyenne de calibration RMSEC et le coefficient de corrélation R^2C , Erreur quadratique moyenne de la validation RMSECV et le coefficient de corrélation de la validation croisée R^2V , et l'erreur quadratique moyenne de prédiction et R^2P de prédiction ont été considérés. De bons modèles devraient avoir RMSEC, RMSECV, RMSEP les plus faibles et les plus élevés R^2 . De plus, les RMSEC, RMSECV, RMSEP ont été utilisés pour déterminer l'optimum modèle sans overfittedness ou underfittedness.

Un résumé des résultats obtenus à partir des différents modèles d'étalonnage à partir de l'huile de tournesol sont donnés dans le Tableau 16.

Tableau 16 : les résultats de différents prétraitements ou combinaisons de traitements faites

| Zone spectrale | Prétraitement | RMSEC | RMSECV | R^2C | R^2V | LV |
|----------------|-----------------------|-------|--------|--------|--------|----|
| 4000-600 | Brut | 0.995 | 1.557 | 0.961 | 0.936 | 6 |
| 3000-1000 | brut | 2.236 | 2.762 | 0.982 | 0.973 | 5 |
| 3000-1000 | centré | 0.838 | 1.174 | 0.986 | 0.974 | 8 |
| 3000-1000 | Centrée normée | 0.780 | 1.344 | 0.988 | 0.966 | 8 |
| 3000-1000 | Dérivée 1(sgolay3-0) | 0.523 | 1.014 | 0.994 | 0.986 | 6 |
| 3000-1000 | snv | 0.961 | 1.342 | 0.982 | 0.967 | 7 |
| 3000-1000 | msc | 0.963 | 1.329 | 0.982 | 0.967 | 7 |
| 3000-1000 | 2dérivée sgolay | 1.061 | 1.408 | 0.978 | 0.963 | 4 |
| 3000-1000 | Lissage sgolay (3-0) | 0.857 | 1.180 | 0.986 | 0.974 | 8 |
| 3000-1000 | Lissage sgolay(3-1) | 0.857 | 1.180 | 0.986 | 0.974 | 8 |
| 3000-1000 | Lissage s golay (5-0) | 0.874 | 1.184 | 0.985 | 0.974 | 8 |

Nous avons fait varier les paramètres suivants :

Les prétraitements mathématiques des spectres : des combinaisons entre les prétraitements, les dérivations du signal et la méthode de régression sont testées (tableau 16).

Après le test de différents prétraitements. Le modèle PLS construit sur la zone spectrale 3000-1000 cm^{-1} pour le tournesol en utilisant le prétraitement de la dérivée première de type Savitzky-Golay a donné les meilleurs résultats de prédiction. Un nombre de 6 variables latentes est obtenu avec une racine carré de la moyenne des carrés de l'erreur de validation de 1.014 RMSEP.

Adultération par l'huile de soja

Un résumé des résultats obtenus à partir des différents modèles d'étalonnage à partir de l'huile soja sont donnés au tableau 17

Tableau 17 : Paramètres de calibration (cross validation) pour le modèle de régression PLS de l'huile de soja

| Zone spectrale | prétraitements | Nombre de Vls | RMSEC % | RMSECV |
|----------------|-------------------------------|---------------|---------|--------|
| 1900-600 | SNV + Smoothing average | 5 | 0.434 | 0.235 |
| 1900-600 | SNV + SMC | 5 | 0.434 | 0.235 |
| 1900-600 | SNV+ Baseline | 5 | 0.42 | 0.22 |

Le modèle PLS construit sur la zone spectrale 1900-700 cm^{-1} pour l'huile de soja à l'aide de Standard Normal Variantes SNV avec un lissage avec le prétraitement SMC a donné le meilleur résultat de prédiction. Un nombre de facteurs égal à 5 pour le modèle de soja était obtenu avec une erreur quadratique minimum pour la validation croisée RMSEP de 0,235.

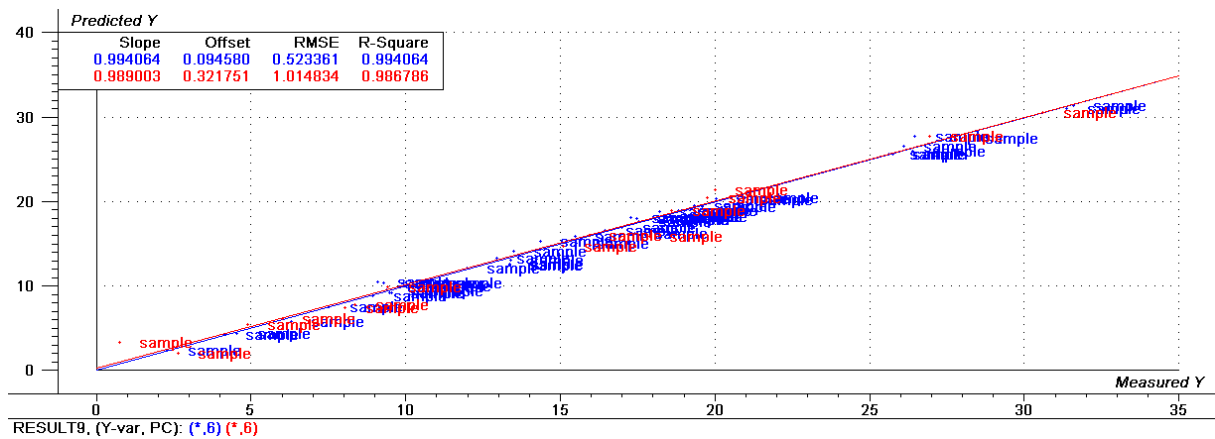


Figure 30a – Les concentrations prédites PLS en fonction des concentrations réelles de l’huile de tournesol dans l’huile d’argan dans l’étape de calibration (bleu) et dans l’étape de validation croisée (rouge) avec la première dérivée de type savitzky- golay

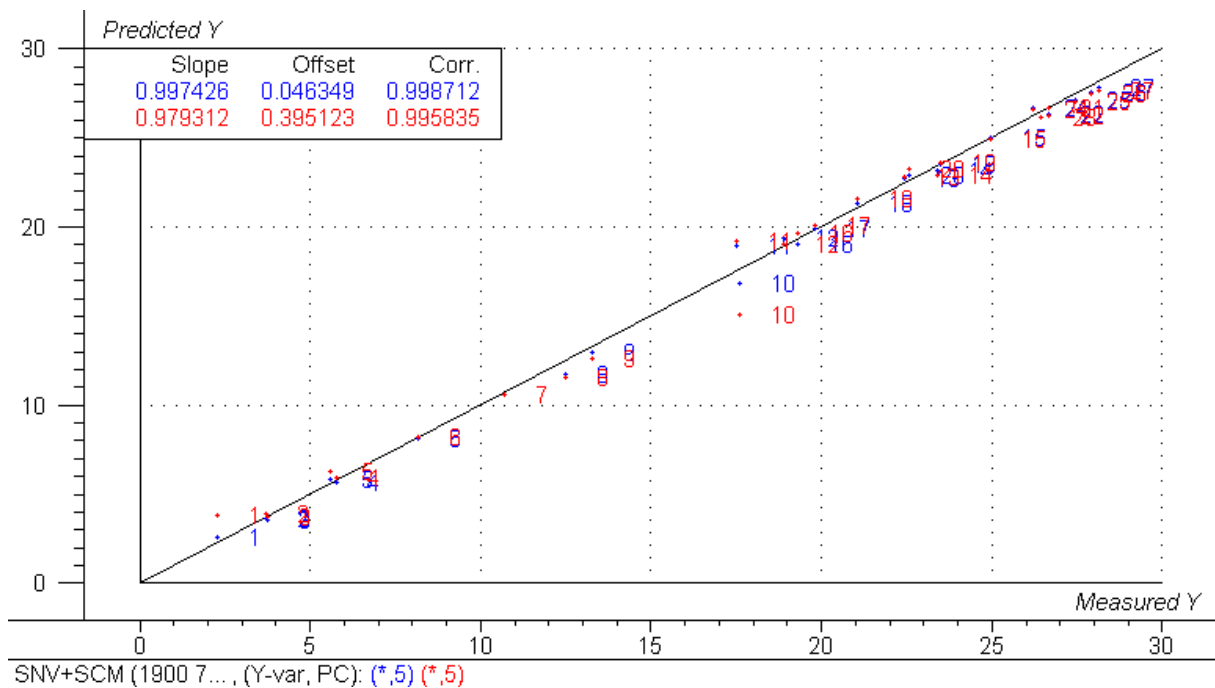


Figure 30b: Les concentrations prédites PLS en fonction des concentrations réelles de l’huile de soja dans l’huile d’argan dans l’étape de calibration (bleu) et dans l’étape de validation croisée (rouge) avec le prétraitement SNV +SCM

Les figures 30a, 30b montrent la concentration réelle par rapport à la concentration prédite en% (p / p) pour la validation croisée et étalonnage.

Les valeurs prédites sont corrélées avec les valeurs mesurées pour le tournesol et le soja à la fois pour un coefficient de corrélation égale à R^2 0,995. De faibles valeurs ont été obtenues pour les deux RMSEC et RMSEP erreur dans

l'étalonnage, qui mesurent l'erreur moyenne dans l'analyse et l'évaluation de la qualité de l'ajustement des données d'étalonnage du modèle développée au cours de l'étalonnage.

Prédiction du modèle

Les résultats de l'huile de tournesol obtenus sur l'ensemble de validation externe sont présentés dans le tableau 18. Cependant, un modèle avec 5 composantes principales a été adopté pour étudier davantage la capacité prédictive pour l'huile falsifiée. Le tableau 19 montre les résultats de prédiction d'un modèle avec 14 échantillons. R carré de 0,990 et 0,834 SEP pour le modèle avec 5 PC a indiqué que ce modèle pourrait être admissible pour la prédiction de l'huile adultérant en comparant ces deux modèles.

Tableau 18 : Résultats de la validation externe pour l'huile de tournesol avec 5 CP

| Concentration réelle | Concentration prédite | Erreur |
|----------------------|-----------------------|--------|
| 4,88 | 5,665 | 0,785 |
| 6,01 | 6,235 | 0,225 |
| 2,63 | 3,031 | 0,401 |
| 8,05 | 7,544 | -0,506 |
| 8,36 | 7,454 | -0,906 |
| 9,42 | 10,133 | 0,713 |
| 15,11 | 14,728 | -0,382 |
| 15,93 | 16,399 | 0,469 |
| 17,88 | 15,861 | -2,019 |
| 18,61 | 18,87 | 0,26 |
| 19,75 | 20,278 | 0,528 |
| 20 | 21,246 | 1,246 |
| 26,94 | 27,628 | 0,688 |
| 30,6 | 30,628 | 0,028 |

Tableau 19 : Les paramètres de prédiction avec le modèle PLS pour le tournesol

| Modèle | Nombre d'échantillons | R-square | Corrélation | SEP | RMSEP | Slope | offset | Bias |
|--------|-----------------------|----------|-------------|-------|-------|-------|--------|-------|
| 5PC | 14 | 0.990 | 0.995 | 0.834 | 0.811 | 1.004 | 0.042 | 0.109 |

Le résultat de la prédiction à partir de soja est représenté dans le tableau 20. La différence entre les concentrations réelles et prédites de l'huile adultérant est faible indiquant la bonne capacité de prédiction du modèle PLS.

Tableau 20: résultat de la validation externe pour l'huile de soja

| Concentration Réelle (%p/p) | Concentration prédite | Error |
|-----------------------------|-----------------------|--------|
| 1.42 | 1.651 | 0.231 |
| 10.69 | 9.406 | -1.284 |
| 16.03 | 15.902 | -0.128 |
| 18.89 | 18.667 | 0.223 |
| 19.35 | 18.79 | -0.56 |
| 23.5 | 22.41 | -1.09 |
| 28.03 | 27.507 | -0.523 |
| 27.7 | 27.273 | -0.427 |

Détection de l'adultération

Le modèle a été en mesure de classer jusqu'à 23,84% (% en poids) le niveau d'adultération de l'huile d'argan par l'huile de tournesol et l'huile de soja.

Tableau 21: le % de la variance pour le modèle ACP

| Principal component Number | Eigen value of Cov(X) | % Variance Captured this PC | % Variance Captured Total |
|----------------------------|-----------------------|-----------------------------|---------------------------|
| 1 | $1.27 e^{-003}$ | 84.20 | 84.20 |
| 2 | $1.27 e^{-004}$ | 7.17 | 91.37 |
| 3 | $7.83 e^{-005}$ | 6.11 | 97.48 |

Le tableau 21 montre qu'avec 3 PC, il est possible d'extraire environ 97% de l'information souhaitée avec RMSEC: 0,00025 et RMSECV: 0,14720 pour la classification entre l'huile d'argan adultérée avec de l'huile de soja et l'huile d'argan adultérée par de l'huile de tournesol. L'analyse ACP a été appliquée aux données FTIR, et il a été constaté que avec trois composantes représentent 98% de la variance totale expliquée ce qui séparé avec succès les mélanges selon la composition.

Conclusion

L'authentification des échantillons d'huile d'argan nécessite généralement l'utilisation de techniques sophistiquées et ne nécessitant pas un long temps d'analyse. Il est nécessaire, pour les tests de contrôle de qualité de disposer de techniques d'analyse simples et rapides.

Cette étude se concentre sur la détection et la quantification de l'adultération de l'huile d'argan par différentes huiles comestibles en utilisant la spectroscopie Moyen-infrarouge (MIR) et la chimiométrie.

Les traitements chimiométriques des spectres MIR ont été évalués pour la classification et la quantification de l'adultération d'argan par l'huile de tournesol ou l'huile de soja. Le potentiel de la spectroscopie infra-rouge combinée avec la régression partielle des

moindres carrés comme technique analytique rapide pour l'analyse quantitative d'adultérants dans l'huile d'argan est démontrée. Un modèle partiel des moindres carrés (PLS) a été créé pour prédire la concentration d'huile de tournesol comme adultérant dans de l'huile d'argan avec une bonne performance de prédiction (R^2 de 0,990 et RMSEP de 0,834). Les bonnes performances de prédiction sont également obtenues pour la prédiction de la concentration de l'huile de soja dans l'huile d'argan.

Les Méthodes de classification multivariées comme l'analyse en composante principale (ACP) peuvent être utilisées. L'ACP a permis de mettre en évidence la différence ou plus particulièrement de classer l'huile de soja adultérant l'huile d'argan à l'huile de tournesol adultérant l'huile d'argan.

Références bibliographiques

- [1] Lai Y. W., Kemsley E. K. , Wilson R. H 1995, Quantitative analysis of potential adulterants of extra virgin olive oil using infrared spectroscopy. *Food Chemistry*, 5, 95-98.
- [2] Mailer, R.J. 2004, *Journal of the American Oil Chemists' Society*,81, 823.
- [3] Hourant ,P.,Baeten,V.,Morales,M.T,Meurens,M.2000 , *Appl.spectrosc*,4 ,1168.
- [4] Wesley ., I.J,Barnes ,R.J,McGill.,A,E,J ,1995, *Journal of the American Oil Chemists' Society*,72,289.
- [5] Yang H., & Irudayaraj J., 2001, Comparison of near-infrared, Fourier transform- infrared, and Fourier transform-Raman methods for determining olive pomace oil adulteration in extra virgin olive oil. *J.of the American Oil Chemists society*, 78, 889-895.
- [6] Schlz,H.,Quilitzsch,R.,Kruger,H 2003, *J.Mol ,Struct*.661:662- 289.
- [7] Lai, Y. W., Kemsley, E. K., & Wilson, R. H. 1994. Potential of Fourier transform infrared spectroscopy for the authentication of vegetable oils. *Journal of Agricultural and Food Chemistry*, 42, 1154–1159.
- [8] Schlz,H.,Ozkan,G.,Baranska,M.,Kruger,H., Ozkan,M. 2005,*vibrat.Spectros.*,39 :249
- [9] Hilali, M., Charrouf, Z., Soulhi, A.E.A., Hachimi, L., Guillaume, D 2007. Detection of argan oil adulteration using quantitative campesterol GC-analysis. *JAOCS, Journal of the American Oil Chemists' Society*, 84 (8), 761-764
- [10] Inon, F. A., Garrigues, J. M., Garrigues, S., Molina, A., & de la Guardia, M. 2003a.Selection of calibration set samples in determination of olive oil acidity by partial least squares-attenuated total reflectance- Fourier transform infrared spectroscopy. *Analytica Chimica Acta*, 489, 59–75.
- [11] Inon, F. A., Garrigues, J. M., Garrigues, S., Molina, A., & De la Guardia, M. 2003b. Selection of calibration set samples in determination of olive oil acidity by partial least squares-attenuated total reflectance-Fourier transform infrared spectroscopy. *Analytica Chimica Acta*, 489, 59–75.
- [12] Savitzky, A., Golay, M.J.E. 1964. Smoothing and differentiation of data by simplified least squares Procedures, *Anal. Chem.* 36, 1627-1639.

CONCLUSION GÉNÉRALE

L'huile d'olive demeure l'un des secteurs stratégiques pour la région méditerranéenne. C'est aussi un secteur en expansion du fait de ses qualités physiologiques de plus en plus reconnues et découvertes par des régions en dehors même de la Méditerranée.

Cette thèse avait des objectifs multiples : le premier était de développer et d'explorer le potentiel des méthodes spectrales comme la spectroscopie moyen infrarouge comme méthode rapide et moins coûteuse pour lier les origines géographiques aux données spectrales (grâce à l'analyse statistique et chimiométrique). Le deuxième objectif était l'évaluation de la qualité des huiles.

L'analyse en composantes hiérarchiques HCA et la classification PLS-DA ont été utilisées afin de construire un modèle de classification pour vérifier l'origine géographique des amandes d'argan à partir de mesures FTIR. L'utilisation combinée de cette technique d'analyse et des méthodes chimiométriques adéquates s'est avérée être une stratégie efficace pour discriminer les amandes d'argan en fonction de leur origine géographique sans avoir besoin d'autres méthodes d'analyse chimique.

Dans ce travail, nous avons montré qu'il est possible d'adapter la spectroscopie infrarouge à transformée de Fourier à la détection de l'adultération de l'huile d'olive par l'huile de soja et l'huile de tournesol. Les modèles reliant ces critères chimiques aux spectres ont été optimisés d'un point de vue chimiométrique. L'influence des méthodes de prétraitement, des méthodes de régression et du domaine spectral sur l'erreur de prédiction ont été étudiées.

L'application de l'analyse en composante principale ACP a permis de faire la distinction entre les huiles d'olive falsifiées par des huiles de soja et de tournesol en utilisant la spectroscopie infrarouge et les méthodes chimiométriques lorsque simultanément deux données de mélanges binaires sont en conjonction dans une matrice de données.

La régression PLS reliant les spectres MIR aux concentrations de l'huile adultérée, a permis d'obtenir une relation satisfaisante entre les concentrations prédites et les concentrations observées. En outre, l'approche PLS est montrée être un outil très important dans la compréhension du processus de calibration du modèle pour l'analyse quantitative des mélanges binaires.

Cette étude a permis de déterminer les acides gras des huiles végétales purs ainsi que les huiles adultérées par chromatographie en phase gazeuse. Nous n'avons pas pu détecter l'adultération de l'huile d'olive par l'huile de soja en se basant sur l'ajout d'un acide gras puisqu'il y a une variation de tous les acides gras avec la variation du taux d'adultération .

En faisant appel à des méthodes chimiométriques comme l'analyse en composantes principales les huiles végétales telles que l'huile d'argan adultérée par l'huile de tournesol sont séparées des huiles d'argan adultérées par l'huile de soja.

La présence d'huile de tournesol et l'huile de soja (comme adultérant) dans l'huile d'argan extra vierge est détectée par la spectroscopie de FTIR en utilisant l'ATR. La capacité de la chimiométrie, comme technique appliquée sur les données MIR-FT a pu classer des huiles sans la nécessité d'une préparation préalable d'échantillons.

Le potentiel de la spectroscopie infra-rouge combinée avec la régression partielle des moindres carrés pour l'analyse quantitative d'adultérants dans l'huile d'argan est démontrée. Un modèle partiel des moindres carrés (PLS) a été créé pour prédire la concentration d'huile de tournesol comme adultérant dans de l'huile d'argan avec une bonne performance de prédiction (R^2 de 0,990 et RMSEP de 0,834). Les bonnes performances de prédiction sont également obtenues pour la prédiction du pourcentage de l'huile de soja dans l'huile d'argan.

Les résultats menés au cours de cette étude nous ont permis également de présenter les méthodes, les hypothèses et les techniques de prétraitement applicables aux spectres Infrarouge. En effet, comme ces spectres contiennent un très grand nombre de variables, les informations fournies ne peuvent pas être extraites directement sans traitement préalable bien adaptée.

Ce travail nous a montré que la spectroscopie infrarouge est un outil très prometteur. Il faut cependant accumuler de nombreuses données expérimentales pour construire des modèles robustes de prédiction de la qualité des huiles d'olive, d'argan ainsi que les amandes d'argan.

Perspectives :

A l'issue des résultats, les actions identifiées et suggérées comme continuation de ce travail se résument dans les points suivants :

- L'étude pourrait s'orienter vers différentes zones du territoire dans la recherche d'une authenticité des huiles
- Pour valoriser davantage ces huiles d'olives et leur donner une valeur ajoutée, l'analyse des autres constituants comme la composition en stérols et en triglycérides s'avère nécessaire.
- Valorisation des sous-produits des huiles
- Etudier l'adultération de l'huile d'argan par l'huile de sésame et l'huile d'arachide
- Application sur d'autres produits alimentaires.

ANNEXE : Liste des publications

1-Publications dans des revues à comité de lecture

1. Abdelkhalek Oussama ,Fatiha Elabadi ,Stefan Platikanov ,Fouzia Kzaiber ,Roma Tauler
Detection of Olive Oil Adulteration Using FT-IR Spectroscopy and PLS with Variable
Importance of Projection (VIP) Scores J Am Oil Chem Soc (2012) 89:1807–1812

3. Fatiha Elabadi, Abdelkhalek.OUSSAMA ,and Olivier.Devos, Analysis of Argan oil
adulteration using infrared spectroscopy .Spectroscopy Letters (2012) vol 45,458-463

2.Fatiha El abadi, Naima Elharfi Abdessamad Balouki ,Fouzia Kzaiber , and
Abdelkhalek.OUSSAMA. Carob origin Classification by FTIR Spectroscopy and
Chemometrics .J chem. :Chem Eng 5(2011)1020-1029

4. Fatiha El abadi , Anna de Juan , Fouzia Kzaiber , Romà Tauler and Abdelkhalek Oussama
.Classification of geographical origin of argan kernels by FTIR spectroscopy, cluster analysis
and partial least squares-discriminant analysis (soumis)

2- Présentations orales

1. Fatiha El abadi , Anna de Juan , Fouzia Kzaiber , Romà Tauler and Abdelkhalek Oussama
A novel chemometric strategy for classification of geographical origin of argan pulpe by
FTIR spectroscopy with cluster analysis and linear discriminant analysis.1^{ère} Rencontre
Internationale de Chimie Moléculaire, Chimiométrie et Applications (RICMCA-2014) Béni-
Mellal le, 29-30 Mai 2014

2.Elabadi Fatiha, Stefan Platikanov, Olivier Devos ,Roma TAULER, Abdelkhalek Oussama,
and Roma Tauler .Detection of oil olive adulteration by sunflower and soya oil using FT-IR
spectroscopy and chemometric.la 4ème Rencontre Internationale sur l'Analyse & la
Chimiométrie (RENACQ-4) 12-13 Mars 2010

3- Présentations par affiche

1.El Abadi Fatiha,Oussama Abdelkhalek ,Kzaiber Fouzia ,Devos Olivier : Study on the
analysis of adulteration in argan oil using infrared spectroscopy and chemometrics. La 4ème
Rencontre Internationale sur l'Analyse & la Chimiométrie (RENACQ-4) 12-13 Mars 2010

2.F.El abadi ,A Oussama,R Tauler: « Application de la spectroscopie IR couplée à la chimométrie pour la détection de l'adultération des huiles », Rencontre Internationale de Spectroscopie et d'optique (RISO 2010) Er-Rachidia, 15 - 16 octobre 2010.

3.F.El abadi ,A Oussama,R Tauler : Detection of olive oil adulteration using FT-IR spectroscopy and chemometric: La 4ème Rencontre Internationale sur la Chimie Moléculaire et Développement, Marrakech, 24-27 Novembre 2010.

4.K Belfkir, F El abadi, A Gorfti ,H Oukani F Kzaiber ,M Mbarki ,A Oussama: Analyse quantitative du taux d'adultération d'huile d'olive par l'huile de tournesol en utilisant la spectroscopie infrarouge FTIR et la chimométrie : La 3ème Rencontre Internationale sur la Chimie Moléculaire et Développement RICMD3, Marrakech, 22-24 Novembre 2007.

Detection of Olive Oil Adulteration Using FT-IR Spectroscopy and PLS with Variable Importance of Projection (VIP) Scores

Abdelkhalek Oussama · Fatiha Elabadi ·
Stefan Platikanov · Fouzia Kzaiber ·
Roma Tauler

Received: 17 January 2012/Revised: 5 May 2012/Accepted: 16 May 2012/Published online: 3 June 2012
© AOCS 2012

Abstract Determination of adulteration and authenticity of extra virgin olive oil (EVOO) was investigated by means of infrared spectroscopy and chemometric methods. The study was focused on the detection and quantification of extra virgin olive oil adulteration by soybean (SB) and sunflower (SF) oils using FT-IR spectroscopy based on the use of PLS modeling and variable importance of projection (VIP) scores. A PLS model, using orthogonal signal correction and mean centering data pretreatments, and VIP scores variable preselection, was able to predict the concentration of sunflower and soybean oil adulterants in the 1–24 % weight ratio range with relative prediction errors lower than 3 % (w/w), for external validation samples. Moreover, the PLS-DA (discriminant analysis) model using the same preselected wavelengths was able to explain 99.9 % of variance and to predict with 100 % accuracy both classes of adulteration (EVOO–SB and EVOO–SF) in the external validation.

Keywords FT-IR spectroscopy · Olive oil · Adulteration · Partial least squares (PLS) · Variable Importance in Projection (VIP) scores

Introduction

Adulteration of food products by the substitution of premium quality ingredients with lower grade and cheaper substitutes [1] is a type of economic fraud that can pose a major health threat to consumers. Extra virgin olive oil (EVOO) shows excellent sensorial and nutritional properties which have been subjected to adulteration by different types of seed oils for long time. Records from the 1980s show that oil adulteration received much attention after causing the death of more than 400 people in Spain and 20,000 casualties after the consumption of adulterated oils [2, 3]. This fact stresses the importance of finding an easy way to detect adulteration, so that the quality of the olive oil used in preparing different foods is of guaranteed quality. There are several different methods used in the detection of oil adulteration. They include the iodine and saponification value methods, the colorimetric reaction based method as well as the refractive index, density and viscosity measurements [4]. Furthermore, UV and fluorescence spectrometry have also been used giving acceptable analytical results [4]. Techniques such as high-performance liquid chromatography (HPLC) and gas chromatography (GC) [5] have also been used for identifying adulterants but these techniques are expensive and time consuming. There have been several attempts to find more rapid and/or efficient methods for the detection of olive oil adulteration. This includes the use of infrared spectrometry both in the mid- and near infrared (IR) regions [6] or the use of the electronic nose systems based on MOS

A. Oussama (✉) · F. Elabadi · F. Kzaiber
Laboratory of Spectroscopy Applied and Environment,
Faculty of Sciences and Technique, Moulay Slimane University,
Beni Mellal, Morocco
e-mail: oussamaabdelkhalek@yahoo.fr

F. Elabadi
e-mail: lscae@fstbm.ac.ma

F. Kzaiber
e-mail: fouziakzaiber@yahoo.fr

S. Platikanov · R. Tauler
Department of Environmental Chemistry, IDAEA-CSIC,
Jordi Girona 18-26, 08034 Barcelona, Spain
e-mail: stefan.platikanov@idaea.csic.es

R. Tauler
e-mail: Roma.Tauler@idaea.csic.es

sensors (HS-E nose) and by solid phase microextraction–gas chromatography/mass spectrometry [7].

Application of IR and Raman spectroscopic techniques combined with chemometric methods is a relatively new approach recently used to determine adulteration of olive oil. The use of these spectroscopic methods for the detection of authentication of vegetable oils first emerged in the middle of 1990s [8, 9]. Near-infrared (NIR), mid-infrared, and Raman spectroscopic techniques were used to quantify the amount of olive–pomace oil adulteration in EVOO (0 and 100 % in 5 % increments by weight) with reasonable results [10] for each technique. Christy [11] used NIR spectroscopy to detect and quantify adulteration of olive oil with soybean, sunflower, corn, walnut and hazelnut oils giving low error intervals. Fourier transform infrared (FT-IR) spectroscopy equipped with a ZnSe-ATR accessory was able to detect adulteration of virgin olive oil with hazelnut oil at levels of 25 % and higher [12]. In the same study, a PLS model was developed to determine the amount of adulteration with a good performance of the proposed method. Vlachos et al. [13] also studied the determination of olive oil adulteration with vegetable oils using FT-IR and giving detection limits for olive oil adulteration of 9 % when the adulterant was corn oil or sesame seed oil and of 6 % when the adulterant was sunflower or soybean oil. Thus, the spectrum of an olive oil sample seems to represent its overall chemical composition and therefore it has the potential to be used for the detection of the olive oil quality [14, 15].

Chemometrics and multivariate data analysis provide help in solving a lot of problems in qualitative and quantitative analysis in general and in adulteration and quality assessment of food products in particular [16]. Adulteration of EVOO by combination with other oils has become a problem for Moroccan olive oil producers and consumers. Specially soybean (SB) and sunflower (SF) oils have been used to adulterate EVOO because of their lower prices in the Moroccan market. SB and SF oils were selected because they are the most frequently used adulterants of EVOO in Morocco markets at present. They are used as adulterants in EVOO because both type of oils are cheap polyunsaturated oils, in which parameters like ω -3 fatty acids and ω -6 fatty acids considered beneficial for human health, are at even higher concentrations.

There are a limited number of published studies focusing on the issue of multiple blended oils used to adulterate EVOO. A chromatographic approach coupled to PLS analysis was proposed for the determination of cottonseed, olive, soybean and sunflower adulteration [17]. The relative standard error of prediction for each oil type in the mixed samples was estimated to be less than 10 %. Recently [18–20] the partial least squares (PLS) and multiple linear regression (MLR) methods were proposed for

the analysis of Fourier transform infrared spectroscopy (FT-IR) spectral data on blended oil samples to estimate qualitatively and quantitatively the presence of vegetable oils (canola, hazelnut, pomace and high linoleic/oleic sunflower, corn, soybean) as adulterants in commercial samples of extra virgin olive oil. Others authors [21] have demonstrated the applicability of mid-IR spectroscopy as a rapid method to detect and quantify the adulteration of EVOO with rapeseed, cottonseed and corn–sunflower binary oil mixtures. The potential adulterants—cottonseed and rapeseed oils—were detected in olive oil at the level of 5 % detection limit in their binary mixtures.

Although in some previous papers the possibility of oil adulteration detection has been already discussed, in most of these experiments only small sampling data sets and short concentration adulteration ranges were employed giving therefore relatively large prediction errors and higher detection limits. Also, the prediction models were built for a specific adulterant and a global model was not considered to predict simultaneously and accurately two types of adulterants in binary mixtures. In contrast, in this paper, we tried to build a calibration model able to predict simultaneously the concentrations of two cheap adulterants in binary mixtures with EVOO, using a significantly larger number of samples and with the aim of achieving lower prediction errors and lower detection limits for the amount of adulteration.

Despite of all the previous chemometrics work done and reported in the literature until now, there is still a need to understand and explore the advantage of chemometrics in oil adulteration problem studies. A considerably amount of effort and costs can be saved if a subset of FT-IR wavenumbers are chosen, based on statistical procedures, with the purpose of identifying key adulteration features. In this work, the methodology we have used and proposed for this purpose is the calculation of PLS-VIP scores.

Here, a new PLS strategy is shown to improve the understanding of the adulteration process and to perform quantitative analysis of binary oil mixtures based on the use and interpretation of the Variable Importance in Projection or VIP scores [22] from the regression model. This is a useful and simple strategy tool for the evaluation of the importance of each variable (spectral wavenumber) in PLS models and it can be used for further spectra wavenumber selection. In summary, the main goal of this study was to test the applicability of mid-IR spectroscopy coupled to chemometrics methods as a means to detect the type and quantity of adulteration of olive oil in binary mixtures at low to high adulteration levels. Also, this study aimed to investigate the performance of the VIP scores for preselecting a reduced number of relevant spectral variables which really have an effect on the response.

Materials and Methods

Samples

Commercial samples of one single origin EVOO, soybean and sunflower vegetable oils were purchased from local grocery stores in Morocco. The binary mixtures EVOO–SB (58 samples) and the EVOO–SF (46 samples) were prepared in the adulteration range of 1–24 % weight ratios.

FT-IR Analysis

The mid infrared (MIR) measurements were acquired with a Bruker Vector 22 FT-IR spectrometer in the spectral region of 4,000–600 cm^{-1} . This instrument was equipped with a attenuated total reflectance (ATR) sampling accessory and a deuterated tri-glycine sulphate (DTGS) detector. The ATR diamond cell was carefully cleaned with pure ethanol to eliminate the presence of oil/fat residues between measurements, i.e. to ensure clean surfaces and obtain the best possible sample spectra. Sample spectra were collected using as a background spectrum, the one obtained with a dry empty ATR cell. One spectrum per sample was recorded. A total number of 104 sample spectra were collected at a resolution of 4 cm^{-1} and 98 scans per sample/background, using OPUS 5.5 software (Bruker Optik GmbH).

Data Analysis

The 104 MIR spectra obtained were imported to Matlab and used to build-up the \mathbf{X} block variables (data matrix \mathbf{X}). Concentrations of the different adulterants (soybean and sunflower oils) in the 104 mixtures were used as the \mathbf{y} block variables. Two data subsets were prepared, one to calibrate the model (with 84 binary mixtures) and another one to validate externally the previous calibration model (with 20 binary mixtures). In both cases, the binary mixtures were randomly selected.

All calculations were performed using PLS Toolbox 5.8 (Eigenvector Research, Manson, WA, USA), and MATA-LAB 6.5 software (The Mathworks, Natick MA, USA).

Chemometrics

The main goal of this study was to build a multivariate model able to explain and predict the changes observed in the concentration of the soybean and sunflower adulterants (\mathbf{y} variables) as a function of the measured MIR spectra (\mathbf{X} block of variables). Partial least squares regression (PLS) and its modification partial least squares with discriminant analysis (PLS-DA) were selected for the modeling of adulterant concentration.

Once experimental data were properly arranged in data matrices, they were mean centered and orthogonal-signal corrected [23–25].

PLSR [26] attempts to maximize the covariance between \mathbf{X} and \mathbf{y} data blocks as it searches for the factor subspace most congruent to both data blocks. A new matrix of weights (reflecting the covariance structure between the \mathbf{X} and \mathbf{y}) is calculated and provided rich factor interpretation information.

In order to perform an authentication study of the EVOO adulteration, PLS-DA, [27] was also used. This technique finds the components or latent variables which discriminate as much as possible between two or more different groups of samples from their FT-IR spectra (\mathbf{X} block) according to their maximum covariance with a target class defined in the \mathbf{y} data block.

The selection of an optimal number of latent variables in PLSR and PLS-DA was done using the criterion of lowest prediction error in cross validation (random subsets) and of optimal prediction of \mathbf{y} values for the external validation samples not used in the calibration step.

The VIP scores of the independent spectral variables (wavenumbers) are a summary of the importance of the projections to find n latent variables. The formula of VIP scores calculation for the j -th variable can be found in [28]. It is important to notice that the average of squared VIP scores equals 1, and the greater than one rule is generally used as a criterion for variable selection [22].

Figures of Merit

The following figures of merit have been calculated to evaluate and validate the different applied methods: root means squared error in calibration (RMSEC), in cross validation (RMSECV) or in the external validation (RMSEP), bias, determination coefficients, LOD (estimated as three times RMSEP). For brevity, in this study the relative prediction errors of concentrations in percentage are included, for both calibration and prediction steps and they are calculated as follows:

$$\text{Rel. error in \%} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i)^2}} \times 100;$$

Quality assessment of the obtained results was discussed by comparison of predicted values versus measured values, both for calibration and validation data sets.

Results and Discussion

Spectral Exploration

Typical spectra between 4,000 and 670 cm^{-1} of the binary mixtures composed by EVOO–SB and EVOO–SF are shown in Fig. 1. The region between 4,000 and 3,035 cm^{-1}

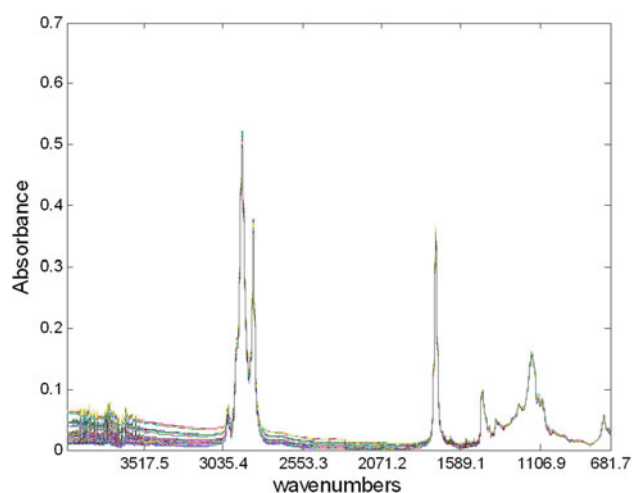


Fig. 1 ATR-MIR spectra of the three binary mixtures in the spectral range 4,000–670 cm^{-1}

was removed because it was due to the water, rich in instrumental noise and providing useless information. [29]. Important spectral bands can be detected in the 2,800–3,035 cm^{-1} wavenumber region because of the C–H stretching mode, in the 1,700–1,800 cm^{-1} region because of C=O stretching and C–O–C stretching and in the wavenumber region of 670–1,400 cm^{-1} because of C–H bending. Many studies based on discriminant analysis have been focused on the entire spectral range or have been limited to the last spectral part of 600–1,400 cm^{-1} .

In this study, the chemometric investigations were performed using as a starting point the spectral region between 3,035 to 670 cm^{-1} of the X data block or in total 2,453 variables.

PLS Partial Least Squares Regression Analysis

Quantification of the relative amounts of adulterant sunflower or soybean oils in the olive oil samples was performed using the PLS regression technique [26]. The investigated range of relative adulterant concentrations in binary mixtures EVOO–SB and EVOO–SF oils was 1–24 % weight ratios of adulteration. In this study, two data pretreatment techniques, mean-centering and OSC [23, 24], were applied. Selection of the lowest number of latent variables was performed by cross validation (leaving one out at time) and by their prediction ability in calibration and validation samples.

The initial PLS investigation was focused on estimating what variables were the more responsible of the achieved results, e.g. what spectral wavenumber regions were more important for optimal prediction of the oil adulteration.

Figure 2 shows the VIP scores obtained for PLS modeling of the binary adulterated mixtures of EVOO–SB and

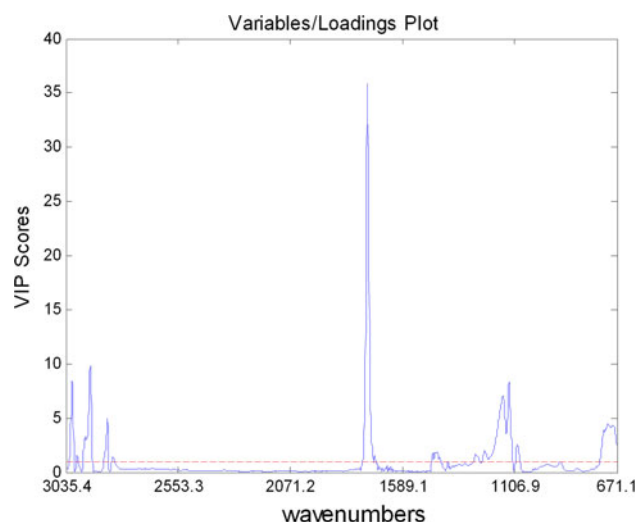


Fig. 2 Variable Importance in Projection (VIP) scores for the investigated spectral region 3,035–671 cm^{-1} of binary adulterated mixtures of EVOO–SB and EVOO–SF oils when they were modeled simultaneously by PLS

EVOO–SF oils in the spectral region 3,035–670 cm^{-1} when mean centering and OSC were applied as pretreatment. Three major spectral zones can be distinguished to be important for optimal PLS model performance. Wavenumbers in the spectral zone between 3,000 and 2,800 cm^{-1} , those around 1,730 cm^{-1} region and a third larger region with many distinctive wavenumbers between 1,450 and 670 cm^{-1} were important. Overall, after the VIP scores variables pre-selection, the data was reduced from 2,453 to 481 variables (wavenumbers), with a total data size reduction, close to 75 %.

Table 1 summarizes the obtained figures of merit for the achieved PLS model built on the reduced data set when only the selected wavenumbers were used.

Figure 3a shows the predicted weight ratios compared to the true weight ratios of soybean and sunflower oils in

Table 1 Figures of merit of the achieved PLS model studying the reduced spectral data after wavenumbers selection with high VIP scores

| Figures of merit | |
|-----------------------------|------------------|
| Model details | OSCMncn and 2LVs |
| Variance explained in % | 99.23 |
| RMSEC (calibration) | 0.57 |
| RMSECV (calibration) | 0.63 |
| RMSEP (external validation) | 0.41 |
| RE % (external validation) | 2.7 |
| R^2 (external validation) | 0.996 |
| Bias (calibration) | 0 |
| Bias (validation) | 0.06 |
| Limit of detection | 1.2 % |

EVOO for calibration samples. This figure shows that a very high correlation between the measured and predicted values of the adulterant in olive oil were obtained (see in Fig. 3a R^2 , RMSEC, RMSECV and bias values). Finally the PLS model yielded a relative prediction error of only 2.7 % in the external validation samples. These figures of merit in both, the calibration and validation steps, confirmed the good predictive abilities of the PLS modeling and confirmed therefore the reliability of the selection method based on the VIP scores for adulteration determinations. However, it should be pointed out that other specific adulteration combinations are possible and they are proposed to be also explored by the proposed PLS-VIP technique. This could be extended for countries with better

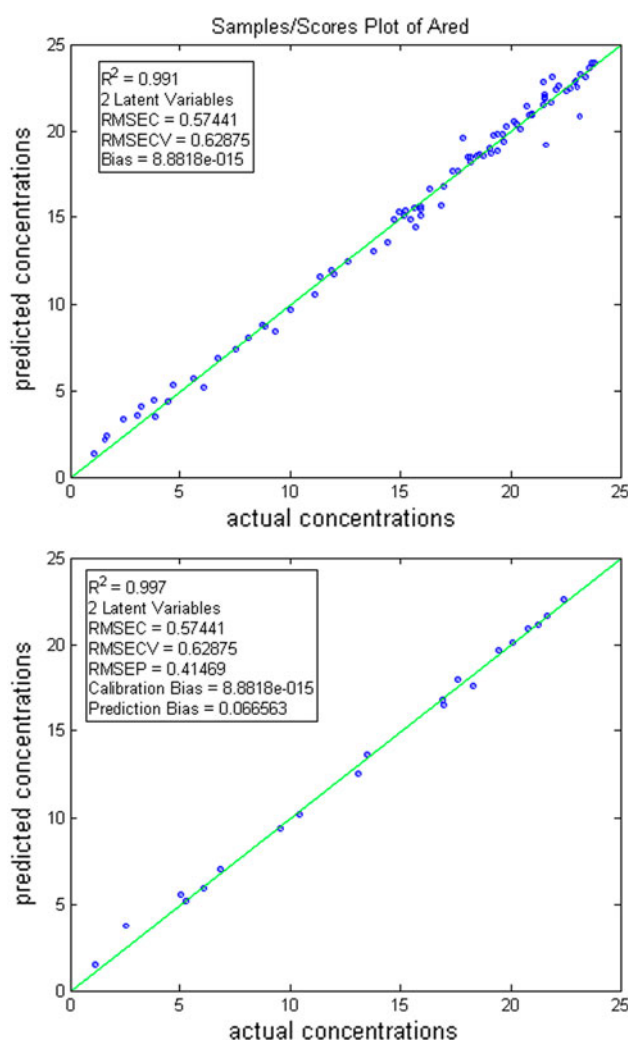


Fig. 3 Predicted versus actual adulteration % (weight ratios) obtained using 2LV PLS model after internal cross validation: **a** calibration; **b** external validation. Reported model figures of merit: LVs, number of latent variables; R^2 , R -square in calibration/validation; RMSEC, root mean squared error in calibration; RMSECV, root mean squared error in internal cross validation; RMSEP, root mean squared error in external validation

established quality controls of different type of market olive oils such as EVOO, virgin, refined or pomace oils. Previous similar studies can be found elsewhere [8, 11], where authors have also proposed the investigation of oils adulteration by mean of other chemometric methods.

PLS-DA, Partial Least Squares Discriminant Analysis Authentication of Adulterants

PLS-DA was applied on the same pre-selected, mean centered and orthogonally signal corrected data of matrix **X** in order to discriminate between EVOO–SF and EVOO–SB mixtures. Two latent variables were found to explain about 99.9 % of the total variance giving a low RMSECV after internal cross validation (random subsets). The first LVs explained more than 99 % of the variance.

Figure 4 illustrates the strong inverse correlation that exists between soybean and sunflower adulterated EVOO samples indicating they produce contrary effects. In this figure, LV1 is shown to discriminate between the two types of oil mixtures.

The PLS-DA model obtained was then applied to the external validation subset of data, consisting of ten adulterated EVOO–SB samples and ten other EVOO–SF samples. FT-IR spectra of these samples were projected on the two LVs subspace and plotted together. The new samples were 100 % correctly identified and classified, showing that the achieved PLS-DA model possesses good predictive ability and that it could be used for future adulteration type authentication analysis and that the pre-selected wavenumbers, based on the VIP scores technique, were also

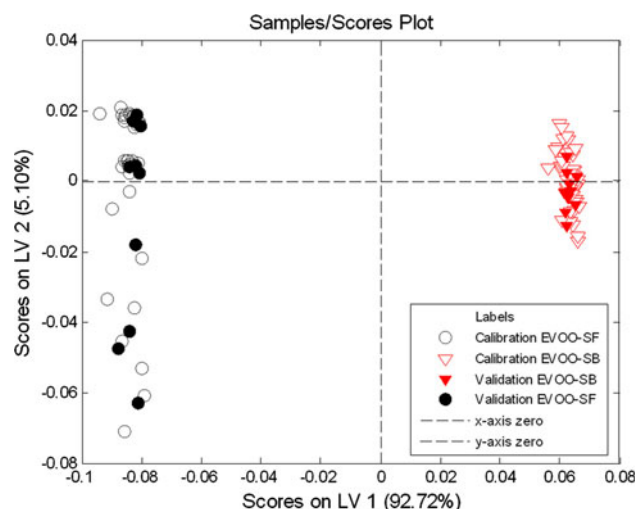


Fig. 4 PLS-DA scores plot (LV1 vs. LV2) in the analysis of the spectral range 3,035–670 cm^{-1} for the adulteration range 1–24 % weight ratios. Samples were binary adulterated mixtures EVOO–SB and EVOO–SF. White/black circles represent EVOO–SF samples used for calibration/validation. White/red triangles represent EVOO–SB samples used for calibration/validation

very informative and influential for discrimination of the class they belong to.

Conclusions

Extra virgin olive oil adulteration with sunflower and soybean oils in binary mixtures can be detected and quantified using FT-IR-ATR spectroscopy coupled with proper chemometric data treatment techniques. PLS and PLS-DA methods are shown to be useful tools for the authentication and further quantification of these adulterants in the concentration range 1–24 % weight ratios. This study revealed that for good adulteration predictions, the wavenumbers pre-selection of the entire spectral region 3,035–670 cm^{-1} was an especially useful practice. The results showed that raw FT-IR spectra could be properly modeled using mean centering and orthogonal signal correction data pretreatment. The PLS model, based on a reduced number of wavelengths selected by VIP scores technique gave good predictions of the adulterant amounts in the external validation, with less than 3 % relative prediction errors. The established PLS-DA model was also capable of authenticating with 100 % accuracy the cheap soybean and sunflower oils as adulterants in extra virgin olive oil at weight ratios as low as 1 %.

References

- Tay A, Singh R, Krishnan S, Gore JP (2002) Authentication of olive oil adulterated with vegetable oils using Fourier transform infrared spectroscopy. *LWT Food Sci Technol* 35:99–103
- Jimeno S (1982) The Spanish toxic symptoms. *Trends Anal Chem* 1:4–6
- Posada M, Castro M, Kilbourne E, Diaz-de Rojas F, Abaitua I, Tabuenca J, Vioque J (1987) Toxic-oil syndrome: case reports associated with the ITH oil refinery in Sevilla. *Food and Chemical Toxicology* 25:87–90
- Tous J (1968) Chemistry and analysis of olive oil. In: Boekennoogen HA (ed) *Analysis and characterisation of oils, fats and fat products*. London: Interscience. Vol 2, 315–606
- Andrikopoulos NK, Giannakis IG, Tzamtzis V (2001) Analysis of olive oil and seed oil triglycerides by capillary gas chromatography as a tool for the detection of the adulteration of olive oil. *J Chrom Sci* 39:137–145
- Guillen M, Cabo N (1997) Infrared spectroscopy in the study of edible oils and fats. *J Sci Food Agric* 75:1–11
- Mildener-Szkudlarz S, Jelen H (2010) Detection of olive oil adulteration with rapeseed and sunflower oils using MOS electronic nose and SPME-MS. *J Food Qual* 33: 21–41
- Baeten V, Meurens M, Morales M, Aparicio R (1996) Detection of virgin olive oil adulteration by Fourier transform Raman spectroscopy. *J Agric Food Chem* 44:2225–2230
- Terouzi W, DeLuca M, Bolli A, Oussama A, Patumi M, Ragno G (2011) A discriminant method for classification of Moroccan olive varieties by using direct FT-IR analysis of the mesocarp section. *Vibr spectr* 56:123–128
- Yang H, Irudayaraj J (2001) Comparison of near-infrared, Fourier transform-infrared, and Fourier transform-Raman methods for determining olive pomace oil adulteration in extra virgin olive oil. *J Am Oil Chem Soc* 78:889–895
- Christy A, Kasemsumran S, Du Y, Ozaki Y (2004) The detection and quantification of adulteration in olive oil by near-infrared spectroscopy and chemometrics. *J Anal Science* 20:935–940
- Ozen B, Mauer L (2002) Detection of hazelnut oil adulteration using FT-IR spectroscopy. *J Agric Food Chem* 50:3898–3901
- Vlachos N, Skopelitis Y, Psaroudaki M, Konstantinidou V, Chatzilazarou A, Tegou E (2006) Application of Fourier transform-infrared spectroscopy to edible oils. *Anal Chim Acta* 573–574:459–465
- Colthup N, Daly L, Wiberly S (1990) *Introduction to Infrared and Raman Spectroscopy*, 3rd edn. Academic Press, Inc., New York
- Skoog D, Holler F, Nieman T (1998) *Principles of Instrumental Analysis*, 5th edn. Saunders College Publishing, Philadelphia
- Muick L, Norgaard L, Englesen S, Bro R, Andersson C (1998) Chemometrics in food science—a demonstration of the feasibility of a highly exploratory, inductive evaluation strategy of fundamental scientific significance. *Chem Intell Lab Sys* 14:31–60
- Hajimahmoodi M, Vander Heyden Y, Sadeghi N, Jannat B, Oveisi M, Shahbazian S (2005) Gas-chromatographic fatty-acid fingerprints and partial least squares modelling as a basis for the simultaneous determination of edible oil mixtures. *Talanta* 66:1108–1116
- Lerma-Garcia MJ, Ramis-Ramos G, Herro-Martinez JM, Simó-Alfonso EF (2010) Authentication of extra virgin olive oils by Fourier-transform infrared spectroscopy. *Food Chem* 118:78–83
- Maggio R, Cerretani L, Chiavaro E, Kaufman T, Bendini A (2010) A new chemometric strategy for the estimation of extra virgin live oil adulteration with edible oils. *Food Control* 21:890–895
- Özdemir D, Öztürk B (2007) Near infrared spectroscopic determination of olive oil adulteration with sunflower and corn oil. *J Food Drug Anal* 15:40–47
- Guardeniz G, Ozen B (1997) Detection of adulteration of extra-virgin olive oil by chemometric analysis of mid-infrared spectral data. *Food Chem* 119:519–525
- Chong I, Chi-hyuck J (2005) Performance of some variable selection methods when multicollinearity is present. *Chem Intell Lab Sys* 78:103–112
- Svensson O, Kourti T, Macgregor J (2002) An investigation of orthogonal signal correction algorithms and their characteristics. *J Chemometrics* 16:176–188
- Wold S, Antic H, Lindgren F, Ohman J (1998) Orthogonal signal correction of near-infrared spectra. *Chem Intell Lab Sys* 44:175–185
- Massart D, Vandeginste B, Buydens L, De Jong S, Lewi P, Smeyers-Verbeke J (1998) *Handbook of chemometrics and qualimetrics*. Elsevier, Amsterdam
- Geladi P, Kowalski B (1986) Partial least-squares regression: a tutorial. *Anal Chim Acta* 185:1–17
- Wold S, Albano C, Dunn WJ, Esbensen K, Hellberg S, Johansson E, Sjostrom M, Martens H, Russwurm J (1983) *Food research and data analysis*. Applied Science, Barking, p 147
- Wold S, Johansson E, Cocchi M (1993) 3D QSAR. In *Drug design; theory, methods, and applications*. ESCOM: Leiden, Holland pp 523–550
- Iñón FA, Garrigues JM, Garrigues S, Molina A, de la Guardia M (2003) Selection of calibration set samples in determination of olive oil acidity by partial least squares-attenuated total reflectance-Fourier transform infrared spectroscopy. *Anal Chim Acta* 489:59–75

Carob Origin Classification by FTIR Spectroscopy and Chemometrics

Fatiha Alabdi, Naima Elharfi, Abdessamad Balouki, Fouzia Kzaiber and Abdelkhalek Oussama*

Département de Chimie et Environnement, Faculté des Sciences et Techniques de Beni mellal, Université Sultan Moulaye Slimane, Beni Mellal 23000, Morocco

Received: April 29, 2011 / Accepted: June 14, 2011 / Published: November 10, 2011.

Abstract: The goal of this study was to use Fourier transform mid-infrared (FTIR) spectroscopy for discrimination of samples of pods and seeds of carob from three Moroccan regions. The origin of samples Pods and seeds of carob could be distinguished from their IR spectra and this measurement was used for discriminate analysis. A multivariate analysis procedure based on the combined use of Hierarchical Cluster Analysis (HCA) and Partial Least Squares-Discriminant Analysis (PLS-DA) was tested and provided good classification results. Three distinctive clusters were recognised, related to the three Moroccan regions. Afterwards, PLS-DA was used for the discrimination and classification of the origin of the various Pods and seeds of carob samples. The results demonstrated that the combined use of FTIR and chemometric analysis (cluster analysis and discrimination by PLS- DA) can be used to rapidly and simply determine the origin of carob pulpe samples.

Key words: FTIR, discriminant analysis, chemometrics, pods of carob, seeds of carob, origin classification, cluster analysis, PLS-DA.

1. Introduction

The carob tree (*Ceratonia siliqua* L.) is an evergreen species with a large distribution around the world and a controversial origin. Zohary [1] suggested that the Mediterranean region has been one of its domestication centres. Carob has been cultivated for thousands of years as a forage crop or food for human consumption [2]. Recently, this species has attracted much attention and became economically important. Pods and seeds are used as raw material in food, pharmaceutical and cosmetic industries [3-5]. It has been introduced and grown in many dry areas of the world. However, its cultivation and production are centered in Spain, Italy and Portugal [6]. Consequently, in the northern side of the Mediterranean basin many cultivars were characterized [5-9]. The characterization was focused

generally on morphological traits.

In Morocco, the production of carob was estimated only to 8% of the world production [10]. This production considered as the fourth range in the world is mainly from natural domesticated trees in agroforestry systems [10]. Indeed, the carob distribution in Morocco is centered in the north selvage of the Atlas chain, the Rif Mountain and in some valleys of the south-west of the Anti-Atlas confined to arid and semi-arid bioclimates with an extension to sub-humid bioclimate in some stands [11-12]. At low altitude, surrounding small aggregate of rural populations, carob trees were empirically selected and propagated. In the agro-forestry systems, carob cultivation remains still traditional and sporadic with no grafting or silvaculture treatment and no fertilization.

The ability of this species to grow in such contrasting climatic and edaphic conditions suggests a high degree of adaptability and a significant variability among its members.

Recently the commercial value of carob has

*Corresponding author: Abdelkhalek Oussama, Professor, research field: environmental chemistry. E-mail: oussamaabdelkhalek@yahoo.fr. Tel: 00210523485112. Fax: 00212523485201.

increased and carob became a plant of multipurpose use [13-17]. In fact, it is used in reforestation program serving both environmental and economic objectives in dry regions, used to valorize marginallands or as substitute for drought sensitive species. Its ecophysiological behaviour has been described as more resistant to water stress compared to other Mediterranean species [18-21].

Noteworthy that if many research reports are available for this species in several areas over the world, little or no data related with its variability and adaptability is available in Morocco. Therefore, it has become imperative to establish a research program aiming at the evaluation of the genetic diversity in the Moroccan carob germplasm.

FTIR spectroscopy is an emergent analytical technique, which presents good sensitivity and a great simplicity in sample preparation.

The infrared spectrum of any compound is known to express a unique “fingerprint”, and this feature enables IR spectroscopy to be used in classification of different samples or identification of unknown samples. Attenuated total reflection (ATR)-Fourier transform infrared (FTIR) spectroscopy techniques have potential quantitative quality control applications for food analyses.

FTIR-ATR provides a simple and reproducible means of handling products in the form of liquids and pastes, the analyses are nondestructive, and the sampling takes less than 2 min. In combination with chemometric techniques, FTIR-ATR methods have been successfully used for the detection of adulteration of fruit purees, honey, and oils [22-24]. Chemometric methods for classification of olive oils have become of increasing importance [25-29] and different Chemometric tools are applied in various laboratories.

This work presents the results of the application of chemometric techniques to the FTIR (Fourier transform infrared spectroscopy) data obtained from samples of Pods and seeds of carob picked in different regions of Morocco. Cluster analysis (CA), principal

component analysis (PCA), and partial least squares regression-discriminant analysis (PLSR-DA) were applied for the characterisation and determination of the origin of the various samples.

2. Materials and Methods

2.1 Samples

The samples were collected from different places in the same region. The collection method, the conditions of storage and preservation are the same for different samples. Investigation was focused on the speeds and pods samples picked in the zones named Fes (fes), Essaouira (ES), Agadir (Aga) and Tafraout (Taf) (Fig. 1). Geographic characteristics such as altitude slice, central latitude and longitude as well as the mean precipitation of these ecoregions are summarized in Table 1. The pulp of carob contains a great majority of its active components. Thus, pulp contains about 7% of proteins, 35% from amidon and 40% sugars. The carob also has a significant calcium, phosphorus rate, of iron, silica, magnesium and pectin. However, the carob contains a low content of greases, rock salt and tanins.

2.2 FT-IR Analysis

A spectrometer Bruker Vector 22 instrument equipped with DTGS detector and Globar (MIR) Source, KBr Germanium separator, was used to measure the spectra of pods and speeds of carob samples. The samples were placed on a single bounce ATR Diamant (Pike). A 98 scans were taken and averaged and measurements were taken at a resolution of 4 cm^{-1} . The background was collected before every sample was measured.

2.3 Chemometric Analysis

Application of the multivariate algorithms was supported by the software The Unscrambler 10.0.1 (Camo Process As., Oslo, Norway). The software elaborate multivariate analysis and experimental design are equipped with several methods, including

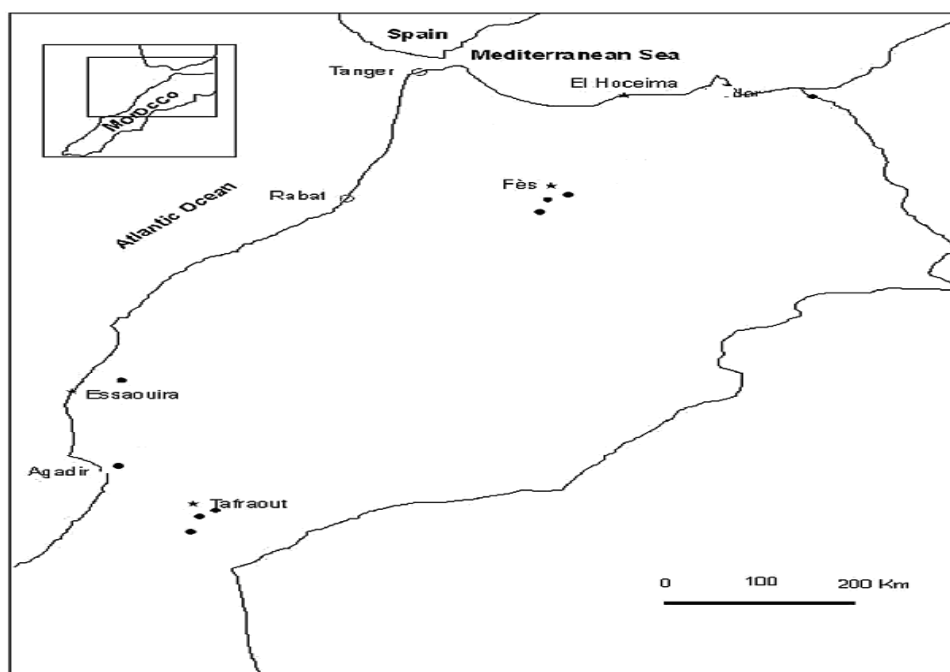


Fig. 1 Map of part of Morocco, “red dots are the 4 area studied”.

Table 1 Clematis and geographical properties of these 4 area.

| Eco-érgions | Code | Latitude N | Longitude w | Altitude (m) | Rainfall (mm) |
|-------------|------|---------------|----------------|-----------------|------------------|
| Essaoira | Ess | 31°20 | 9°40 | 100-200 | 300 |
| Tafraout | Taf | 29°42 | 8°59 | 1000-1200 | 190 |
| Fes | Fes | 33°30 | 4°56 | 1100-1300 | 600 |
| Agadir | Aga | 30°41 | 9°33 | 150-350 | 300 |

CA, PCA, regression methods, and PLS-DA. It also allows to optimise the calibration models and to develop validation procedures.

3. Results and Discussion

3.1 Pods of Carob

FT-IR spectra of samples, cultivated and produced in different locations of Morocco were recorded and divided in two sets: a calibration set of 32 samples and a prediction set of 12 samples. Fig. 2 shows the FTIR spectra of the studied pods of carob. The differences between them were clearly small and manifest just in limited zones of the spectra. The use of single peaks or selected wavelength ranges to obtain information about the botanical origin of the pods of carob is very hard. These data were so conveniently handled by multivariate statistical techniques.

With the aim to obtain more information from the FTIR spectral data, the spectra were firstly subjected to mathematical elaboration.

In particular, Multiplicative Scatter Correction (MSC), Standard Normal Variate (SNV) and Derivative elaborations were applied [30-31]. The best improvement in data variance was reached when the derivative function through the Savitzky-Golay algorithm was used. Best results were obtained by fixing the following parameters: 1 order, number of smoothing points 3 and polynomial order 2.

3.1.1 Principal Component Analysis

Fig. 3a shows the PC1/PC2 score plot, in which the first two PCs summarised more variation in the data than any other PCs, accounting for 96.71% of the data variance for pods carob with 5 PC. All the classes resulted perfectly separated from the other ones. The Fig. 3b shows the evolution of the variance explained by the number of principal components. From 4 CP obtains the maximum information on the model.

3.1.2 Cluster Analysis for Pods Carob

The calibration set obtained from derivative transformation of the FT-IR data was employed to perform the CA, applying the Square Euclidean Distance.

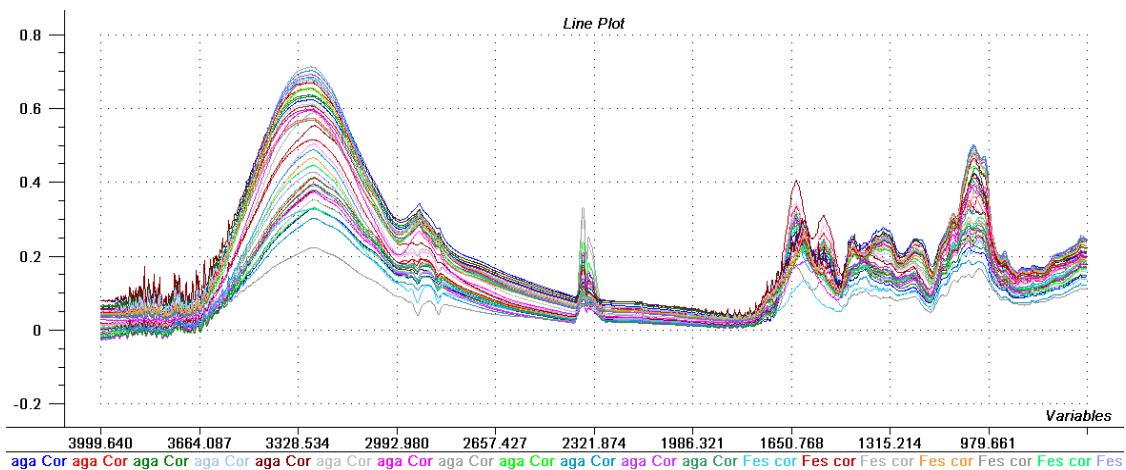


Fig. 2 FT-IR spectra of calibration samples for pods of carob.

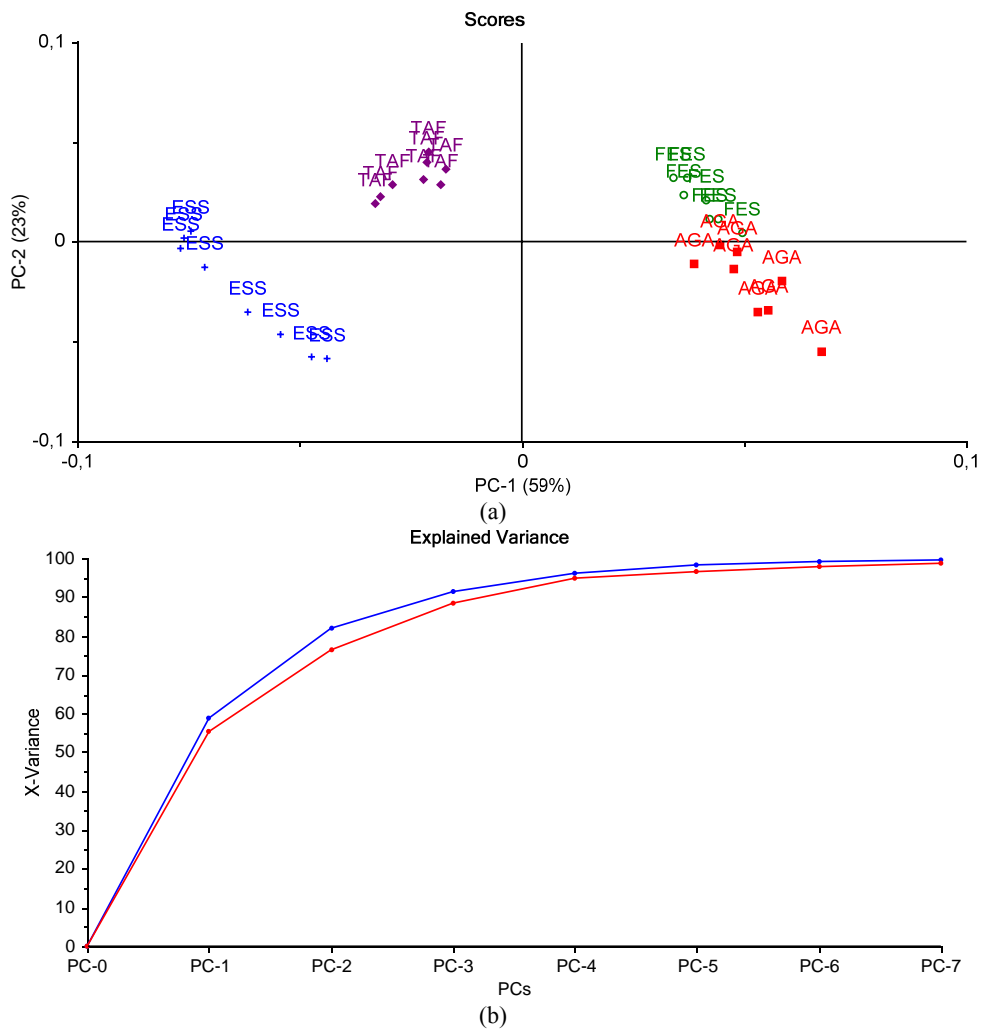


Fig. 3 (a) Score plot of pods carob for PCA 1D3.2 SG with 5 pc; (b) Explained variance with 5 pc of 1D SG prtratament PCA 96.71%.

On the basis of the connecting distances and in agreement with the PCA results, results were reported

in the form of dendrogram, shown in Fig. 4. Four distinctive clusters were defined.

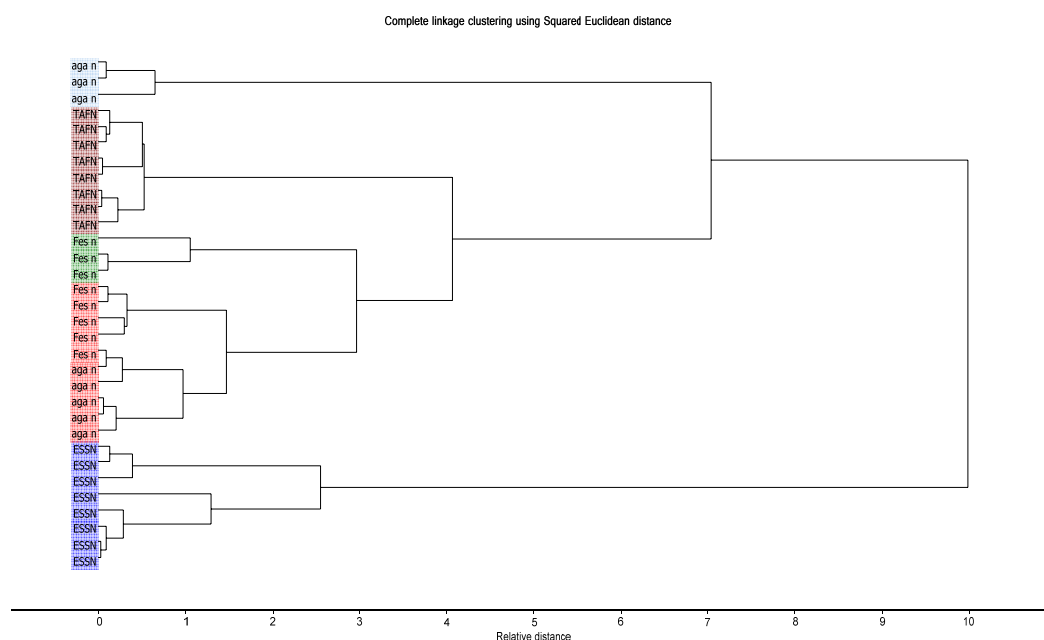


Fig. 4 HCA dendrogram of Agadir, Fes, Essaouira and Taфраout pods carob using complete linkage squares euclidean distances.

The first group contained three objects and was assorted with samples collected in a geographical area of Agadir. The second cluster was formed by eight samples, all localized in the territory Taфраout. The third cluster was formed by eight numbers localized in Fes. The fourth cluster was formed by five objects localized in area of Agadir. The last cluster contained eight samples located in Essaouira area. These indicate that the four origins form natural clusters except only one of Agadir which divide into two groups.

3.1.3 Classification Model

In the classification step, a Partial Least Squares Discriminant Analysis (PLS-DA) was performed to classify speeds and pods samples of the calibration set, but picked in different times. The procedure used a PLS2 algorithm, fixing a Y-variable for each Class from 1 to 4.

The classification model was validated by Full Cross Validation and the statistical results are listed in Table 2 for pods carob. The correlation coefficient (R^2), ranging between 0.7978 and 0.9443, and the root mean square error of prediction (RMSEP), from 0.2108 to 0.4019 for pods of carob could be considered satisfactory.

Table 2 Statistical parameters carried out by full cross validation on PLS-DA model (pods of carob).

| PCs = 7 | Classes | | | |
|---------|---------|--------|--------|--------|
| | Agadir | Fes | Ess | Taf |
| R^2 | 0.7978 | 0.8029 | 0.9443 | 0.8467 |
| RMSEP | 0.4019 | 0.3967 | 0.2108 | 0.3499 |

In order to show the differences among the pods carob regions, the score plot using the first factor and the second factor was used (Fig. 5). Four regions (group) can be clearly identified. Each group represents one type of pods. This result clearly shows the ability of this technique to discriminate among the four pods carob used in this study.

A further optimization procedure was tempted by a careful selection of the wavelength regions to be used in calibration to ensure an acquisition of the most useful information and a removal of redundant zones or noisy data. The Martens' Uncertainty Test was chosen as the tool to select the most useful wavelengths. This approach was unable to separate the information contained in the full spectra and to identify the fingerprint for each class of different seeds and pods of carob.

3.1.4 Classification of New Samples.

The final classification model was applied on the

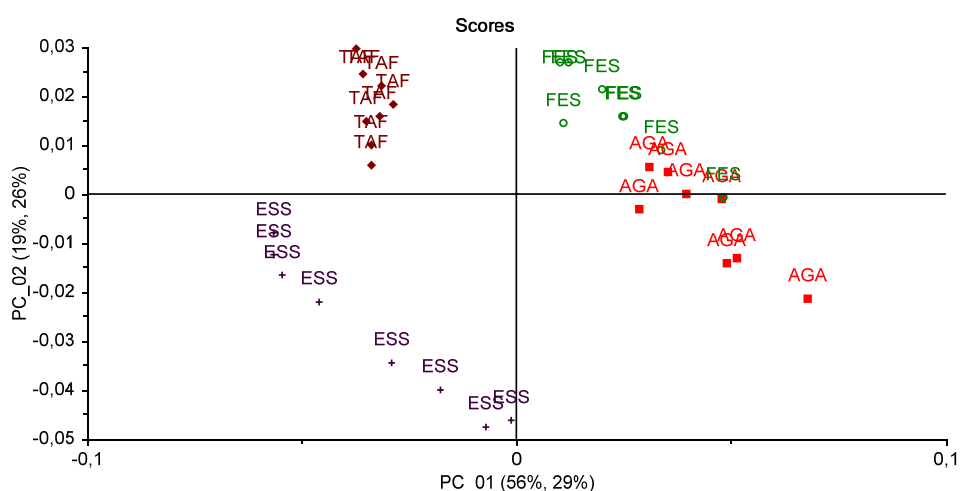


Fig. 5 Score plot of PLS-DA model for pods of caroubier of Ess, Agadir, Fes and Tafilat.

prediction set composed of the FTIR data from analysis of twelve (12) new pods samples. This operation allowed classifying the new objects into the classes established with the previous classification built on the calibration set. A sample was considered belonging to a class when X was superior value of 0.5.

Table 3 shows the classification results for pods carob with the comparison between the predicted values of each Y and the theoretical reference classes.

3.2 Seeds of Carob

With the aim to obtain more information from the FTIR spectral data, the spectra were firstly subjected to mathematical elaboration.

In particular, Multiplicative Scatter Correction (MSC), Standard Normal Variate (SNV) and Derivative elaborations were applied [30-31]. The best improvement in data variance was reached when the Multiplicative Scatter Correction (MSC) was used.

Principal component analysis

Fig 6 shows the superposition of the infrared spectra of different samples, it shows that the maximum information is contained in the area of 700 to 1400 cm^{-1} with less noise.

Fig.7a is a dendrogram for different samples, it is clear that the samples are grouped into families according to their similitude.

Fig. 7b shows that 4PC explain 97.17% of variance

total for speeds of carob.

3.2.1 Cluster Analysis for Speeds of Carob

The calibration set obtained from Multiplicative Scatter correction transformation of the FT-IR data was employed to perform the CA, applying the Euclidean Distance. On the basis of the connecting distances and in agreement with the PCA results, Results were reported in the form of dendrogram, shown in Fig. 8. four distinctive clusters were defined.

The first group contained four objects and was assorted with samples collected in a geographical area of Fes. The second cluster was formed by eight samples, all localized in the territory of Agadir. The third cluster was formed by eight objects all localized in Essaouira. The last cluster contained eight samples located in Tafrouat. This indicates that the four origins form natural clusters.

3.2.2 Classification Model

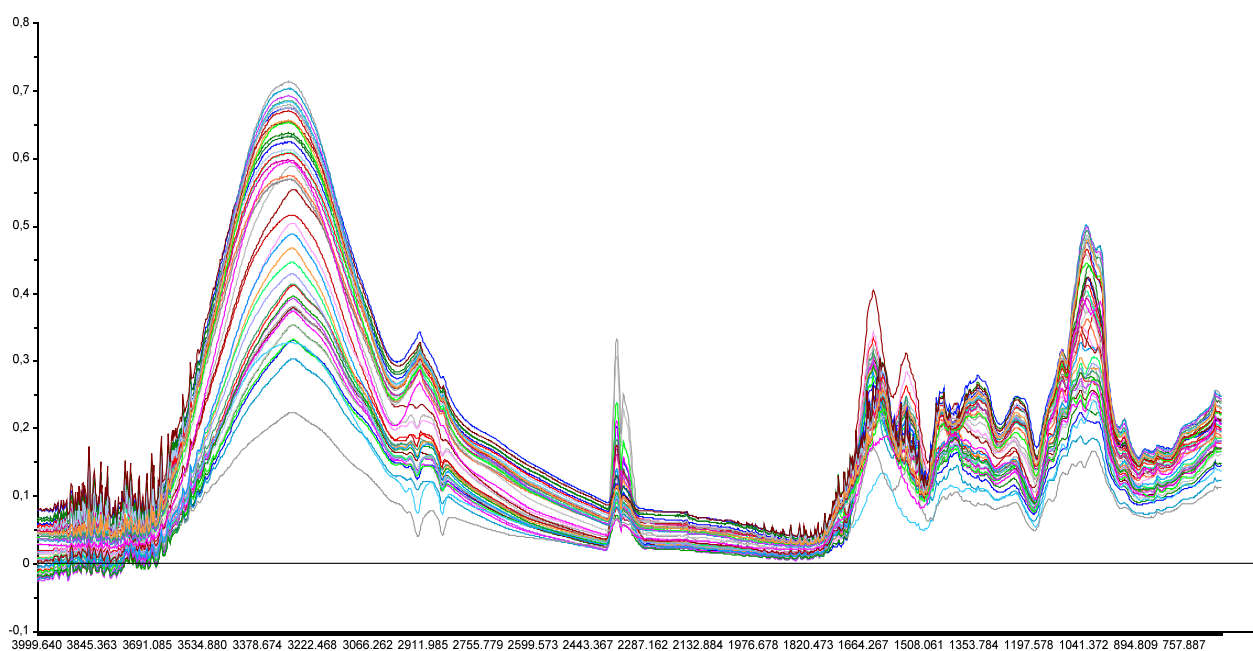
In the classification step, a Partial Least Squares Discriminant Analysis (PLS-DA) was performed to classify speeds samples the calibration set, but picked in different times. The procedure used a PLS2 algorithm, fixing a Y -variable for each Class from 1 to 4.

The classification model was validated by Full Cross Validation and the statistical results are listed in Table 5 for seeds of carob.

The correlation coefficient (R^2), ranging between 0.7073 and 0.9372, and the root mean square error of

Table 3 classification of new samples of pods of carob.

| | Agadir | | Fes | | Ess | | Taf | |
|-----|------------|-----------|------------|-----------|------------|-----------|------------|-----------|
| | Prediction | Reference | Prediction | Reference | Prediction | Reference | Prediction | Reference |
| Ag | -0.0845 | 1 | -0.2020 | -1 | -0.7476 | -1 | -0.9659 | -1 |
| Ag | -0.0941 | 1 | -0.2089 | -1 | -0.7419 | -1 | -0.9551 | -1 |
| Ag | 0.0659 | 1 | -0.0941 | -1 | -0.8373 | -1 | -1.1346 | -1 |
| Ag | -0.0346 | 1 | -0.1661 | -1 | -0.7774 | -1 | -1.0219 | -1 |
| Fes | -0.1110 | -1 | -0.2210 | 1 | -0.7318 | -1 | -0.9362 | -1 |
| Fes | -0.1078 | -1 | -0.2187 | 1 | -0.7337 | -1 | -0.9398 | -1 |
| Fes | -0.1382 | -1 | -0.2405 | 1 | -0.7156 | -1 | -0.9057 | -1 |
| Fes | -0.2572 | -1 | -0.3259 | 1 | -0.6447 | -1 | -0.7722 | -1 |
| Ess | -1.2622 | -1 | -1.0467 | -1 | -0.0458 | 1 | 0.3547 | -1 |
| Ess | -1.3148 | -1 | -1.0845 | -1 | -0.0144 | 1 | 0.4137 | -1 |
| Ess | -1.2974 | -1 | -1.0720 | -1 | -0.0248 | 1 | 0.3941 | -1 |
| Ess | -1.2192 | -1 | -1.0159 | -1 | -0.0714 | 1 | 0.3064 | -1 |
| Taf | -0.9805 | -1 | -0.8447 | -1 | -0.2136 | -1 | 0.0388 | 1 |
| Taf | -0.9847 | -1 | -0.8477 | -1 | -0.2111 | -1 | 0.0435 | 1 |
| Taf | -1.0693 | -1 | -0.9083 | -1 | -0.1607 | -1 | 0.1383 | 1 |
| Taf | -1.1414 | -1 | -0.9601 | -1 | -0.1178 | -1 | 0.2192 | 1 |

**Fig. 6** FT-IR spectra of calibration samples for seeds of carob.

prediction (RMSEP), from 0.2428 to 0.4675 for speeds of carob could be considered satisfactory (Table 4).

In order to show the differences among the seeds carob regions, the score plot using the first factor and the second factor was used (Fig. 9). Four regions (group) can be clearly identified. Each group represents one type of seeds carob. This result clearly shows the

Table 4 Statistical parameters carried out by full cross validation on PLS-DA model (seeds of caroub).

| PCs = 5 | classes | | | |
|----------------|---------|--------|--------|--------|
| | Agadir | Fes | Ess | Taf |
| R ² | 0.8217 | 0.7073 | 0.9372 | 0.8830 |
| RMSEP | 0.3760 | 0.4675 | 0.2488 | 0.2936 |

ability of this technique to discriminate among the four speeds and pods used in this study.

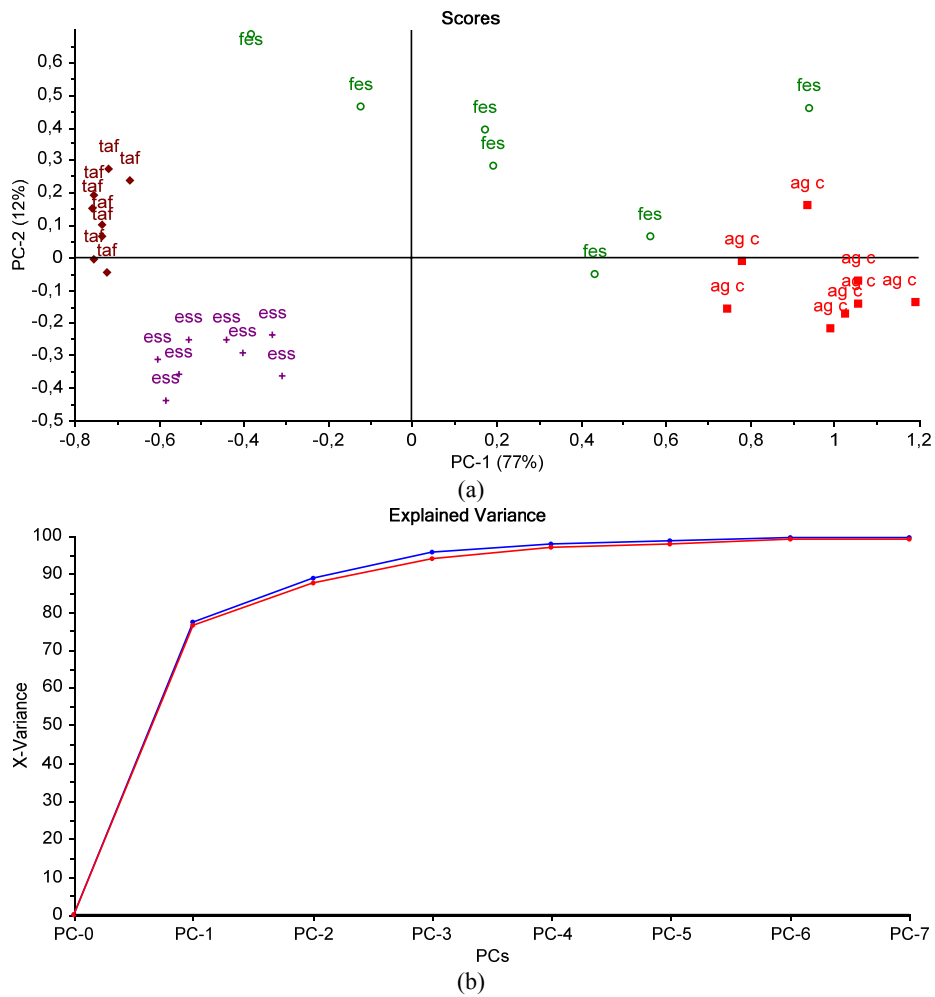


Fig. 7 (a) Score plot of seeds of carob for PCA (MSC) with 4 pc; (b) Explained variance with 4 pc of Msc pretreatment PCA 97.17%.

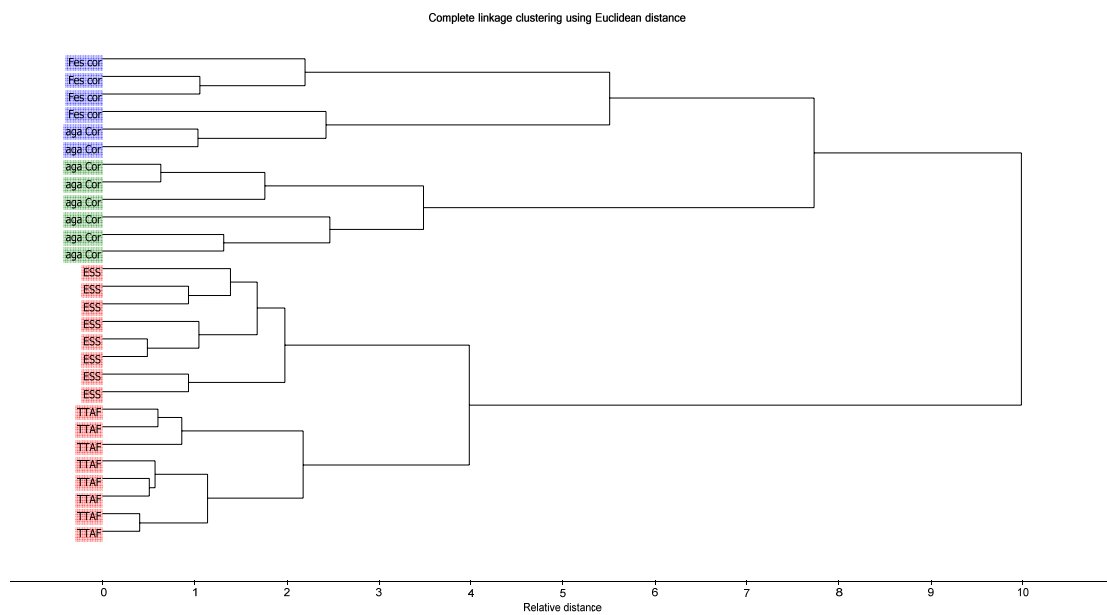


Fig. 8 HCA dendrogram of Agadir, Fes, Essaouira and Taфраут of seeds carob using complete linkage euclidean distance.

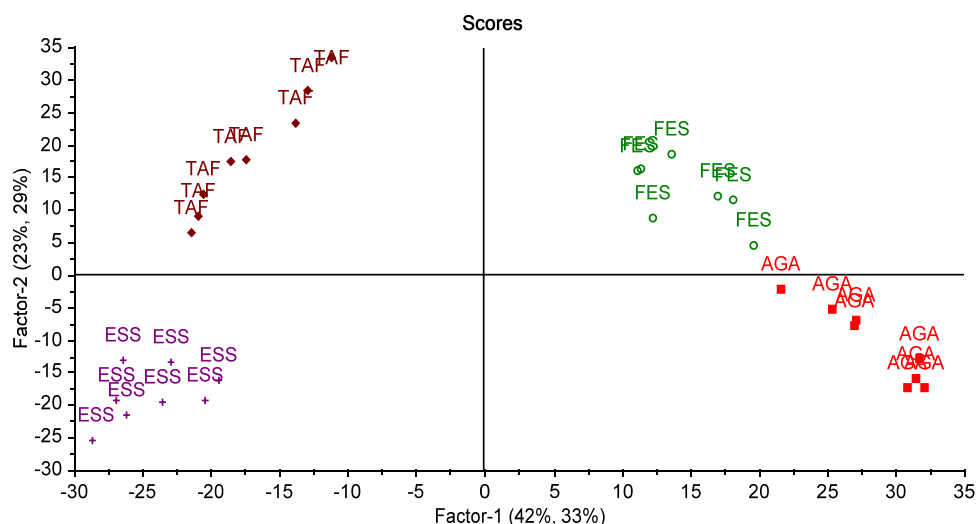


Fig. 9 Score plot of PLS-DA model for seeds of caroubier of Ess, Agadir, Fes and Tafilat.

Table 5 Classification of new samples of seeds of carob.

| | Agadir | | Fes | | Ess | | Taf | |
|-----|------------|-----------|------------|-----------|------------|-----------|------------|-----------|
| | Prediction | Reference | Prediction | Reference | Prediction | Reference | Prediction | Reference |
| Ag | 0.4739 | 1 | -0.4583 | -1 | -1.1383 | -1 | -0.8772 | -1 |
| Ag | 0.4972 | 1 | -0.5981 | -1 | -0.8886 | -1 | -1.0105 | -1 |
| Ag | 0.8663 | 1 | -0.6867 | -1 | -0.8941 | -1 | -1.2856 | -1 |
| Ag | 0.5090 | 1 | -0.5974 | -1 | -0.9370 | -1 | -0.9745 | -1 |
| Fes | 0.2106 | -1 | -0.3781 | 1 | -1.0754 | -1 | -0.7572 | -1 |
| Fes | -0.7732 | -1 | 0.4778 | 1 | -0.9579 | -1 | -0.7468 | -1 |
| Fes | -0.5485 | -1 | 0.3634 | 1 | -1.0403 | -1 | -0.7746 | -1 |
| Ess | -0.9723 | -1 | -1.0545 | -1 | 1.1044 | 1 | -1.0176 | -1 |
| Ess | -0.7434 | -1 | -1.1510 | -1 | 0.8167 | 1 | -0.9224 | -1 |
| Ess | -1.1246 | -1 | -0.9920 | -1 | 1.1288 | 1 | -1.0123 | -1 |
| Ess | -1.0454 | -1 | -1.0384 | -1 | 0.9412 | 1 | -0.8574 | -1 |
| Taf | -0.8294 | -1 | -1.1714 | -1 | -0.5451 | -1 | 0.5459 | 1 |
| Taf | -0.8826 | -1 | -1.1069 | -1 | -0.8596 | -1 | 0.8491 | 1 |
| Taf | -0.8863 | -1 | -1.1538 | -1 | -0.6360 | -1 | 0.6761 | 1 |

4. Conclusions

A convenient method to classify seeds and pods of carob from different regions without any chemical separation is proposed. The classification procedure is based on the use of FTIR data that confirmed their utility to be used as finger prints of the sample. A series of seeds and pods from Morocco were separated into four classes. Classification was then used to predict the origin of new samples from unknown origin. The analytical differences in the FTIR spectra were amplified by transforming the original data by derivative function for pods carob and by multiplicative

scatter correction for seeds carob. The successful application of the model proved the ability of the chemometric classification methods applied on IFTR data in the characterization of seeds and pods of carob.

References

- [1] Zohary, M. *Geobotanical Foundations of the Middle East*. Vols 2. Stuttgart, 1973.
- [2] Zohary, M. Domestication of the Carob (*Ceratonia Siliqua* L). *Israel J. Plant Sci.* **2002**, *50*, 141-145.
- [3] Batista, M. T.; Amaral, M. T.; Proenca Da Cunha, A. Carob Fruits as Source of Natural Oxidant. In *Third International carob Symposium*, Proceedings of the Communication, Tavira, Portugal, June 19-23, 1996.

- [4] Vourdoubas, J.; Makris, P.; Kefalas, J.; Kaliakatsos, J. Studies on the Production of Bioethanol from Carob. In *Industry and Climate Protection*, Proceedings of the 12th National Conference and Technology Exhibition on Biomass for Energy, Amsterdam, 2002; pp 489-493.
- [5] Barracosa, P.; Osorio, J.; Cravador, A. Evaluation of Fruit and Seed Diversity and Characterization of Carob (*Ceratonia siliqua* L) Cultivars in Algarve Region. *Sci. Hortic.* **2007**, *114*, 250-257.
- [6] Batlle, I.; Tous, J. Cultivares Autoctonos de Algarrobo (*Ceratonia siliqua* L.) en Catalunya. *Invest. Agr.* **1990**, *5(2)*, 223-238.
- [7] Orphanos, P. I.; Papaconstantinou, J. The Carob Varieties of Cyprus. *Tech. Bull. Cyprus Agricultural Research Institute*, Nicosia, 1969, 5.
- [8] Albanell, E.; Caja, G.; Plaixats, J. Characterization of Carob Fruits (*Ceratonia siliqua* L.), Cultivated in Spain for Agroindustrial Use. *Int. Tree Crops J.* **1996**, *9*, 1-9.
- [9] Russo, G.; Polignano, G. B. Variation of Seed and Fruit Characters in *Ceratonia siliqua* L. Cultivars. *Genet. Resour. Crop Evol.* **1996**, *43*, 525-553.
- [10] Batlle, I.; Tous, J. Promoting the Conservation and Use of Underutilised and Neglected Crops 17 Carob Tree *Ceratonia siliqua* L. *Institute of Plant Genetics and Crop Plant Research*, Gatersleben/International Plant Genetic Resources Institute, Rome, Italy, 1997, p 92,.
- [11] Aafi, A. Contribution à l'étude Phytoécologique et à La Cartographie des Groupements Végétaux du Parc National de Talassemtane. Mém. de 3^{ème} cycle, ENFI, Salé, 1995, p 162, plus annexes.
- [12] Emberger, L.; Maire, R. *Catalogue des Plantes du Maroc (Spermatophytes et Pte'ridophytes)*. Imprimerie Minerva, Alger, 1941.
- [13] Carlson, W. A. The Carob: Evaluation of Trees, Pods and Kernels. *Int. Tree Crops J.* **1986**, *3*, 281-290.
- [14] Roukas, T. Continuous Ethanol Productions from Carob Pod Extract by Immobilized *Saccharomyces Cerevisiae* in a Packed Bed Reactor. *J. Chem. Technol. Biotechnol.* **1994**, *59*, 387-393.
- [15] Corsi, L.; Avallone, R.; Cosenza, F.; Farina, F.; Baraldi, C.; Baraldi, M. Antiproliferative Effects of *Ceratonia Siliqua* L, on Mouse Hepatocellular Carcinoma Cell Line. *Fitoterapia* **2002**, *73*, 674-684.
- [16] Makris, D.; Kefalas, P. Carob Pods (*Ceratonia siliqua* L) as a Source of Polyphenolic antioxidants. *Food Technol. Biotechnol.* **2004**, *42(2)*, 105-108.
- [17] Sandolo, C.; Coviello, T.; Matricardi, P.; Alhaique, F. Characterization of Polysaccharide Hydrogels for Modified Drug Delivery. *Eur. Biophys. J.* **2007**, *36(7)*, 693-700.
- [18] Winer, N. The Potential of the Carob (*Ceratonia siliqua* L). *Int. Tree Crops J.* **1980**, *1*, 15-26.
- [19] Nunes, M. A.; Rauralho, J. D. C.; Rijo, P. S. Seasonal Changes in Some Photosynthetic Properties of *Ceratonia siliqua* (Carob-Tree) Leaves under Natural Conditions. *Physiol. Plant.* **1992**, *86*, 381-387.
- [20] Rejeb, M. N. Etude des mé'Canismes de re' Sistance a` la se' Cheresse du Caroubier. *Rev. Res. Ame' lior. Prod. Milieu. Aride.* **1992**, *1*, 47-55.
- [21] Sakcali, M. S.; Ozturk, M. Eco-Physiological Behaviour of Some Mediterranean Plants as Suitable Candidates for Reclamation of Degraded Areas. *J. Arid Environ.* **2004**, *57*, 1-13.
- [22] Irudayaraj, J.; Xu, F.; Tewari, J. Rapid Determination of Invert Cane Sugar Adulteration in Honey Using FTIR Spectroscopy and Multivariate Analysis. *Journal of Food Science* **2003**, *68*, 2040-2045.
- [23] Ozen, B. F.; Weiss, I.; Mauer, L. J. Dietary Supplement Oil Classification and Detection of Adulteration Using Fourier Transform Infrared Spectroscopy. *Journal of Agricultural and Food Chemistry* **2003**, *51*, 5871-5876.
- [24] Tay, A.; Singh, R. K.; Khrisnan, S. S.; Gore, J. P. Authentication of Olive Oil Adulterated with Vegetable Oils Using Fourier Transform Infrared Spectroscopy. *LWT Food Science and Technology* **2002**, *35*, 99-103.
- [25] Aranda, F.; Gómez-Alonso, S.; Rivera del Álamo, R. M.; Salvador, M. D.; Fregapane, G. Triglyceride, Total and 2-Position Fatty Acid Composition of Cornicabra Virgin Olive Oil: Comparison with Other Spanish Cultivars. *Food Chemistry* **2004**, *86(4)*, 485-492.
- [26] Brescia, M. A.; Alviti, G.; Liuzzi, V.; Sacco, A. Chemometric Classification of Olive Cultivars Based on Compositional Data of Oils. *Journal of the American Oil Chemists' Society* **2003**, *80(10)*, 945-950
- [27] Casas, J. J. S.; Bueno, E. O.; Garcia, A. M. M.; Cano, M. M. Study of Fatty Acid Contents in Monovariety Oils Elaborated from Olives Produced in Extremadura (Spain). *Grasas y Aceites* **2003**, *54(4)*, 371-377.
- [28] Cichelli, A.; Pertesana, G. P. High-Performance Liquid Chromatographic Analysis of Chlorophylls, Pheophytins and Carotenoids in Virgin Olive Oils: Chemometric Approach to Variety Classification. *Journal of Chromatography A* **2004**, *1046(1-2)*, 141-146.
- [29] Luca, M. D.; Terouzi, W.; Ioele, G.; Kzaiber, F.; Oussama, A.; Oliverio, F.; Tauler, R.; Ragno, G. *Food Chem.* 2011, doi:10.1016/j.foodchem.2010.07.010.
- [30] Iñón, F. A.; Garrigues, J. M.; Garrigues, S.; Molina, A.; de la Guardia, M. Selection of Calibration Set Samples in Determination of Olive Oil Acidity by Partial Least Squares-attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy. *Analytica Chimica Acta* **2003**, *489*, 59-75.
- [31] Maggio, R. M.; Kaufman, T. S.; Del Carlo, M.; Cerretani, L.; Bendini, A.; Cichelli, A.; et al. Monitoring of Fatty Acid Composition in Virgin Olive Oil by Fourier Transformed Infrared Spectroscopy Coupled with Partial Least Squares. *Food Chemistry* **2009**, *114*, 1549-1554.



Analysis of Argan oil adulteration using FTIR infrared spectroscopy and chemiometria

| | |
|-------------------------------|---|
| Journal: | <i>Journal of the American Oil Chemists Society</i> |
| Manuscript ID: | JAOCs-10-0501 |
| Manuscript Type: | Letter to the Editor |
| Date Submitted by the Author: | 25-Oct-2010 |
| Complete List of Authors: | abdelkhalek, oussama; Faculte des sciences and technics, chemistry and environment fatiha, elabadi; fst beni mellal, chemistry |
| Keywords: | Lipid Chemistry / Lipid Analysis, Oleochemistry, Spectroscopy < Lipid Chemistry / Lipid Analysis |
| | |

SCHOLARONE™
Manuscripts

view

Analysis of Argan oil adulteration using FTIR infrared spectroscopy and chemiometria

Fatiha Elbadi¹, Olivier.Devos², and Abdelkhalek.OUSSAMA¹

¹ Laboratoire de Spectrochimie Appliqué et Environnement, Faculté des Sciences et Techniques de Béni Mellal, Université Moulay Slimane, Maroc, oussamaabdelkhalek@yahoo.fr

² Université des Sciences et Technologies de Lille (USTL), Laboratoire de Spectrochimie Infrarouge et Raman (LASIR), olivier.devos@univ-lille1.fr

Abstract

The authentication of argan oil samples requires usually the use of sophisticated and time consuming analytical techniques. There is a need for fast and simple analytical techniques quality control testing.

The determination of argan adulteration by vegetables oils is a real analytic challenge. This study focuses on the detection and quantification of argan oil adulteration with different edible oils using mid-infrared (IR) spectroscopy with chemometrics. Chemometric treatment of MIR spectra was assessed for the classification and quantification of argan oil adulteration with sunflower or soya oil. The potential of infrared spectroscopy combined with partial least squares regression as a rapid analytical technique for the quantitative determination of adulterants in Argan oil is demonstrated. A partial least squares (PLS) model was established to predict the concentration of sunflower oil as an adulterant in argan oil with good prediction performance (R^2 of 0.990 and RMSEP of 0.834). Good prediction performance is also obtained for the prediction of soya oil concentration in argan oil. Furthermore multivariate classification methods like Principal Component Analysis (PCA) can be used to classify soya oil adulterated argan oil to sunflower oil adulterated argan oil.

Keywords: Argan oil, Infrared spectroscopy, chemiometria, adulteration.

Introduction

The determination of food authenticity and the detection of adulteration are problems of increasing importance in the food industry. This is especially true for 'value-added' products, where the potential financial rewards for substitution with a cheaper ingredient are high. The consumers are not in a position to detect these inexpensive components in the food to which he /she pays a high price for.

Argan oil is obtained from the Argan fruit which are present only in south of Morocco. Argan oil obtained from the Argan tree (*Argania spinosa*) represents about 25 % of dietary fats consumed by the southwestern Moroccan population. It is characterized by its unique fatty acid composition. This oil is

1
2
3 particularly rich in oleic (45% to 48%) and linoleic (32% to 35%) acids and quite devoid of
4 polyunsaturated fatty acids of the w-3 family [1]. It is currently used in traditional medicine for its
5 cosmetic, bactericide and fungicide properties. This oil receives increasing attention due to its
6 pharmacological potential like the reduction in cardiovascular risk [2] and the presence of anticancer
7 substances [3]. Argan oil is produced in low yield and its preparation is time consuming.
8 Consequently, argan oil is expensive. In 2005, in developed countries the average price of a bottle of
9 100 mL was around 20 US\$. Such a price may incite people to substitute part of argan oil by cheaper
10 edible oil and therefore the control of quality of argan oil is an important issue.

11 Most of the methods used to detect the adulteration of oil are based on liquid and gas chromatography
12 [4]. However, all these methods are time-consuming and require skilled operators. More recently,
13 Fourier transform infrared (FTIR) spectroscopy has been shown to have potential for discriminating
14 between extra virgin and other oils [5] The chemical variation between extra virgin and other oils is
15 such that very high spectral quality is required, imposing rigorous experimental demands, in order to
16 successfully apply discriminant analysis. However, Lai et al [5] were able to develop a successful
17 discriminant model and provide a method suitable for use by an untrained user, with an analysis time
18 of approximately 2-3 min.

19 Molecular spectroscopy techniques such as mid-infrared (MID-IR) or near infrared (NIR) have been
20 shown to be useful for quality control of olive oils [6-7] for example the rapid evaluation of
21 adulteration of virgin

22 olive oils [8-9]. For qualitative analysis, the mid (MID-IR) spectroscopy has more application because
23 the fingerprints of functional groups can be displayed narrowly and intensely in the MIR region (4000-
24 400 cm^{-1}). Fourier transform infrared (FTIR) spectroscopy with attenuated total reflectance (ATR) or
25 transmission cell accessories has been used to authenticate, identify or classify fats and oils [10-12].

26 So far, few studies are available on the discrimination of Argan oil from other edible oil [13] and most
27 of them were carried out by gas chromatography. The main objective of this paper is to demonstrate
28 the potential of infrared spectroscopy for the quantitative determination of the levels of adulterant in
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 argan oil by sunflower and soya oil. The method uses FTIR spectroscopy, partial least squares (PLS)
4
5 multivariate analysis and principal component analysis.
6
7

8 **Material and methods**

9 **Samples**

10
11
12
13
14
15 The Argan, sunflower and soya oils were obtained from a local grocery. The argan oil was adulterated
16
17 with sunflower oil at levels between 0 and 30% (w/w) or soya oil at levels between 0 and 30% (w/w).
18
19 A total of 67,37 samples were prepared, with 53,29 assigned for use as a calibration set for method
20
21 development, and 14,8 as a validation set for sunflower and soya oils respectively.
22
23

24 **Instrumentation**

25
26
27 All spectra were collected on a Bruker vector 22 FTIR spectrometer, equipped with deuterated
28
29 triglycine sulphate detector and Globar source. Measurements were performed by using 128 scans at 4
30
31 cm^{-1} resolution between 4000 and 500 cm^{-1} . Samples were deposited, without preparation on an
32
33 Attenuated Total Reflectance (ATR) cell provided with a diamond crystal. The background spectrum
34
35 was recorded on air for each sample.
36
37

38 **Statistical analysis**

39
40
41
42 PLS analysis was accomplished after transferring the data into the UNSCRAMBLER 9.6 package
43
44 (CAMO, Norway). Different Pre-treatment combining standardization (mean centering and scaling to
45
46 unit variance all spectral data points), derivatives... have been tested. The optimum number of factors
47
48 was selected using the process of internal cross validation.
49
50

51 **Results and discussion**

52 **Adulteration of Argan oil**

53
54
55
56
57
58
59 FTIR spectra of the Argan oil, sunflower oil and soya oil are reported in the figure 1.
60

1
2
3 Figure 2 shows the spectra of Argan with different level of sunflower oil in the 3200-600 cm^{-1} and
4
5 Figure 3 the spectra of Argan with different level of soya oil in the 700-1900 cm^{-1} . It can be noticed
6
7 than no characteristic peak for one of the 2 oils is observed in this spectral region.

8
9
10 Furthermore for the PLS analysis, the region between 2400 and 2260 cm^{-1} was deleted because of its
11
12 low signal to noise ratio and the presence of fluctuation that is independent of the sample composition
13
14 [14-15]. The region between 4000 and 3030 cm^{-1} was also removed because it contained no useful
15
16 chemical information and contributed to instrumental noise.

17
18
19 The quantification of the sunflower or soya oil contents of the adulterated argan oil samples was
20
21 performed using PLS1 regression. Different spectra pre-processing has been tested and different
22
23 spectral zone selected prior PLS analysis.
24

25 26 **Result of cross validation**

27
28
29 **Derivative:** like smoothing, this transformation is relevant for variables which are themselves a
30
31 function of some underlying variable (e.g. absorbance at various wavelengths). The *Savitzky-Golay*
32
33 algorithm was chosen for these studies [16]. The calibration models were developed using PLS
34
35 regression with full cross validation. All of the samples for modelling was split into a calibration set
36
37 with full cross validation (n=53) and predict set (n=14) for sunflower samples and with (n=37) for
38
39 soya full cross validation and (n=8) for predict set. To evaluate the results, the Root Mean Square
40
41 Error of calibration and correlation coefficient (r) for the calibration set, root mean square error of
42
43 cross validation and r of cross validation, and root mean square error of prediction and r of prediction
44
45 were considered. Good models should have lower RMSEC, RMSECV, RMSEP and higher r. In
46
47 addition, the minimum plot of the RMSEC, RMSECV, RMSEP were used to determine the optimal
48
49 model without "overfittedness" or "underfittedness" [17].
50
51
52

53
54 A summary of the results obtained from the different calibration models from sunflower are given in
55
56 Table 1.
57

58
59 The PLS model constructed on the 3000-1000 cm^{-1} spectral zone for sunflower using first derivative of
60
savitzky-Golay pretreatments yielded to the best prediction results. The curve showing the evolution

1
2
3 of the RMSEP according to the number of factors shows that the cross validation error is lower for 6
4
5 factors with 1.014 RMSEP.
6
7

8 A summary of the results obtained from the different calibration models from soya are given table 2.
9
10 The PLS model constructed on the zone $1900-700\text{ cm}^{-1}$ spectral zone for soya using standard normal
11
12 variate SNV with a smoothing SMC pretreatments yielded to the best prediction results. The number of
13
14 factors from soya shows that the cross validation error is minimum for 5 factors with 0.235 RMSEP.
15
16

17
18 Figures 3a, 3b shows the actual versus predicted concentration in % (w/w) for cross validation and
19
20 calibration. The values predicted are correlated with the measured values at R 0.995 for both
21
22 sunflower and soya.
23
24

25 Low values were obtained for both RMSEC and RMSEP error in calibration, which measure the
26
27 average error in the analysis and evaluate the goodness of fit of the calibration data to the model
28
29 developed during calibration.
30
31

32 **Result for extern validation**

33
34
35 The sunflower results obtained on the validation set are presented in table 3. However, a
36
37 model with 5 principal components was adopted to further investigate the predictive ability
38
39 for adulterated oil. Table 4 shows the prediction results from a model with other 14 samples.
40
41 R square of 0.990 and SEP of 0.834 for model with 5 PC indicated that this model could be
42
43 eligible for prediction of the adulterant oil comparing these two models. The prediction result
44
45 from soya is shown in a table 5. Difference between the actual and predicted adulterant oil
46
47 concentration is small indicating the good prediction ability of PLS model.
48
49
50
51
52

53 **Adulteration detection**

54
55
56 The model was able to classify up to 23.84 % (weight %) adulteration level of sunflower and soya oil.
57
58 Table-1 showed that with 3 PC it is possible to extract about to 97% of desired information with
59
60

1
2
3 RMSEC: 0.00025 and RMSECV: 0.14720 to classification between oil olive adulterate with soya and
4
5 olive oil adulterate with sunflower oil.
6
7

8 In the first stage of the analysis PCA was applied to the processed FTIR data, and it was found that
9
10 Three PCs accounting for 98% of the total explained variance successfully separate the mixtures
11
12 according to composition.
13
14

15 16 Conclusion

17
18 It can be concluded that adulteration of argan oil can be monitored using FTIR spectroscopy with the
19
20 ATR sampling technique. The potential of infrared spectroscopy combined with partial least squares
21
22 regression as a rapid analytical technique for the quantitative determination of adulterants in Argan oil
23
24 is demonstrated. Partial least squares can be successfully used to detect the level of adulterant, such
25
26 sunflower and soya oil. A partial least squares (PLS) model was established to predict the
27
28 concentration of sunflower oil as an adulterant in Argan oil with good prediction performance R-
29
30 square of 0.990 and RMSEP of 0.834.
31
32
33

34
35 A partial least squares (PLS) model was established to predict the concentration of soya oil as an
36
37 adulterant in argan oil with good prediction performance.
38
39

40 Further work is needed to take into account the spectral variability of Argan oils and to test and extend
41
42 this method to other edible adulterant oil.
43
44

45 Multivariate classification methods like Principal Component Analysis (PCA) can be used to
46
47 classify soya oil adulterated argan oil to sunflower oil adulterated argan oil.
48
49
50

51 52 References

- 53
54
55 [1] Berrougui H, Ettaib A, Herrera Gonzalez MD, Alvarez de sotomayer(2004). Hypolipidemic and
56
57 hypocholesterolemic effect of argan oil (*Argania spinosa*) lowers blood pressure and improves endothelial
58
59 dysfunction in spontaneously hypertensive rats. *Br J Nutr*, 92: 921-9,
60

- 1
2
3 [2] Keys A.: Mediterranean diet and public health: personal reflections (1995). *Am. J Clin Nutr*, 61 (suppl.):
4 1321-1323,
5
6
7
8 [3] Khalloukhi F., Younos C., Soulimani R. Oster T., Charrouf Z., Spiegelhalter B., Bartsch H., Owen R.W
9 (2003), Consumption of argan oil with its unique profile of fatty acids, tocopherols, squalene, sterols and
10 phenolic compounds should confer valuable cancer chemopreventive effects. *Eur J Cancer Prev*, 12: 67-75,
11
12
13
14
15
16 [4] Mondello, L., Casilli, A., Tranchida, P.Q., Costa, R., Chiofalo, B., Dugo, P., Dugo, G (2004), Evaluation of
17 fast gas chromatography and gas chromatography-mass spectrometry in the analysis of lipids. *Journal of*
18 *Chromatography A*, 1035 (2): 237-247
19
20
21
22
23
24 [5] Lai Y. W., Kemsley E. K. , Wilson R. H (1995), Quantitative analysis of potential adulterants of extra virgin
25 olive oil using infrared spectroscopy. *Food Chemistry*, 5: 95-98.
26
27
28 [6] Mailer, R.J. (2004) , *Journal of the American Oil Chemists' Society*, 81: 823.
29
30
31
32
33 [7] Hourant ,P.,Baeten,V.,Morales,M.T,Meurens,M.(2000) , *Appl.spectrosc*,4 : 1168.
34
35
36
37 [8] Wesley ., I.J,Barnes ,R.J,McGill.,A,E,J(1995), *Journal of the American Oil Chemists' Society*,72: 289.
38
39
40
41 [9] Yang ,H.,Irudayaraj,J(2001), *Journal of the American Oil Chemists' Society*,78:889.
42
43
44
45 [10] Schlz,H.,Quilitzsch,R.,Kruger,H (2003), *J.Mol ,Struct*.661:662- 289.
46
47
48
49
50 [11] Lai ,Y.W.,Kemsley,E.K,Wilson,R (1994) *Journal of Agricultural and Food Chemistry*,42 :1154.
51
52
53
54 [12] Schlz,H.,Ozkan,G.,Baranska,M.,Kruger,H., Ozkan,M. (2005),*vibrat.Spectros*39 :249
55
56
57
58
59
60

- 1
2
3 [13] Hilali, M., Charrouf, Z., Soulhi, A.E.A., Hachimi, L., Guillaume, D (2007) Detection of argan oil
4 adulteration using quantitative campesterol GC-analysis. *JAOCS, Journal of the American Oil Chemists'*
5 *Society*, 84 (8): 761-764
6
7
8
9
10
11 [14] Inon, F. A., Garrigues, J. M., Garrigues, S., Molina, A., & de la Guardia, M. (2003a). Selection of
12 calibration set samples in determination of olive oil acidity by partial least squares-attenuated total reflectance-
13 Fourier transform infrared spectroscopy. *Analytica Chimica Acta*, 489: 59–75.
14
15
16
17
18
19 [15] Inon, F. A., Garrigues, J. M., Garrigues, S., Molina, A., & De la Guardia, M. (2003b).
20 Selection of calibration set samples in determination of olive oil acidity by partial least squares-attenuated total
21 reflectance-Fourier transform infrared spectroscopy. *Analytica Chimica Acta*, 489: 59–75.
22
23
24
25
26
27 [16] Savitzky, A., Golay, M.J.E. (1964) Smoothing and differentiation of data by simplified least squares
28 Procedures, *Anal. Chem.* 36: 1627-1639.
29
30
31
32
33 [17] Naes, T., Fearn, T., & Davies, A.M. (2002). A user-friendly guide to multivariate calibration and
34 classification. UK : NIR Publication .
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

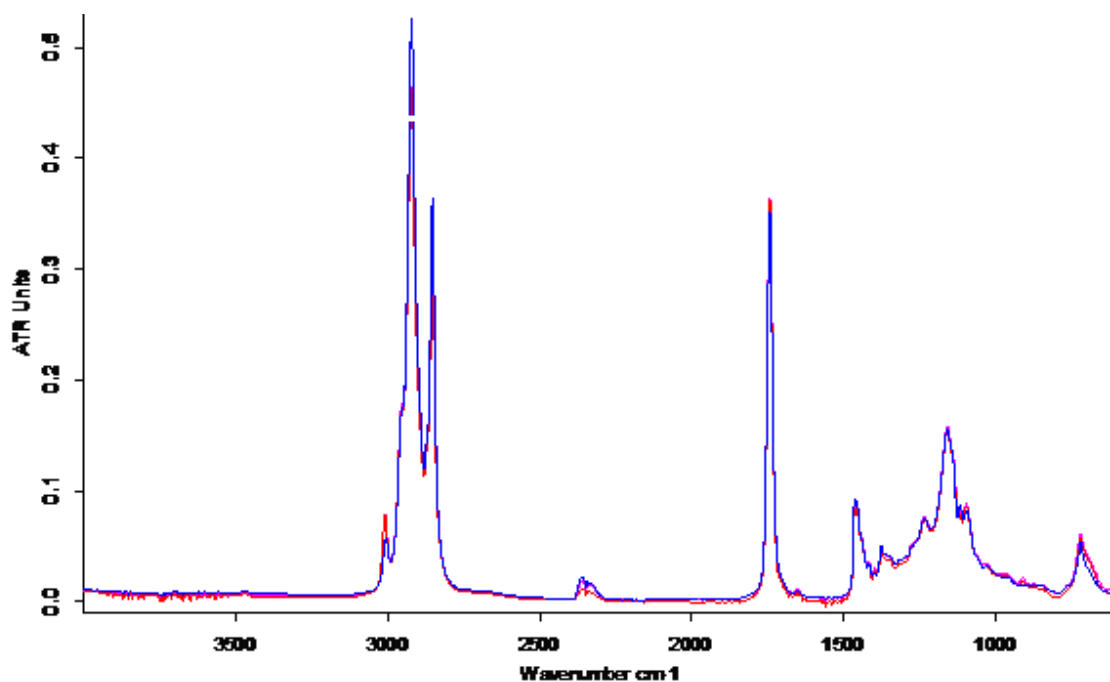


Figure 1 – spectra FTIR of argan oil (blue), soya oil (red) and sunflower oil (rose)

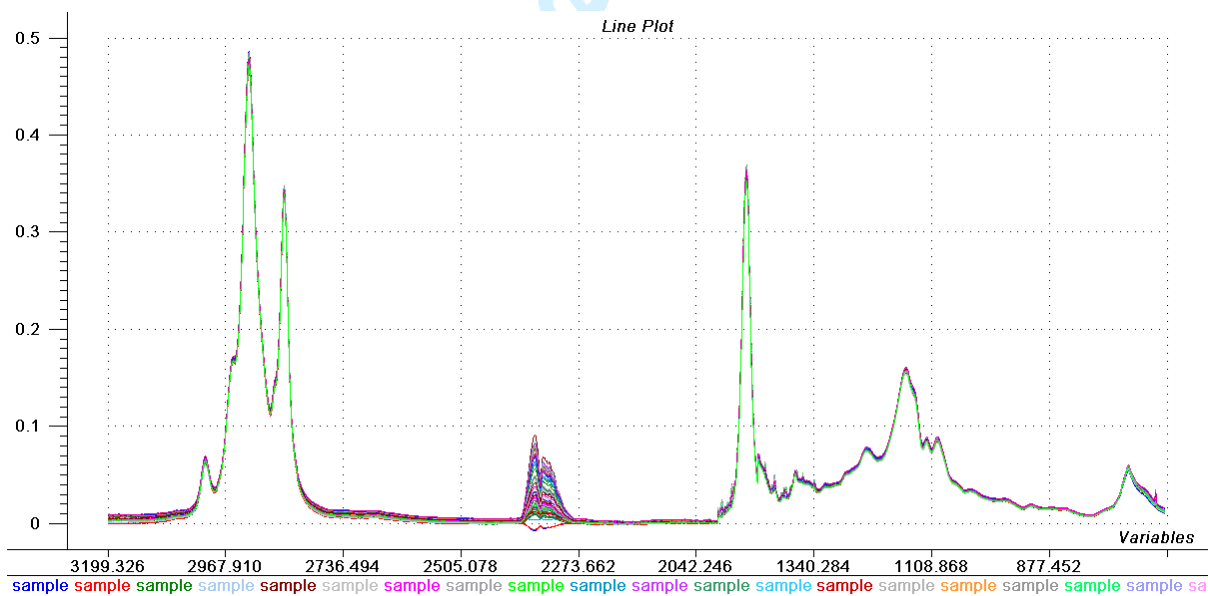


Figure 2 – spectra of argan-sunflower oil samples in 3200-600 cm⁻¹

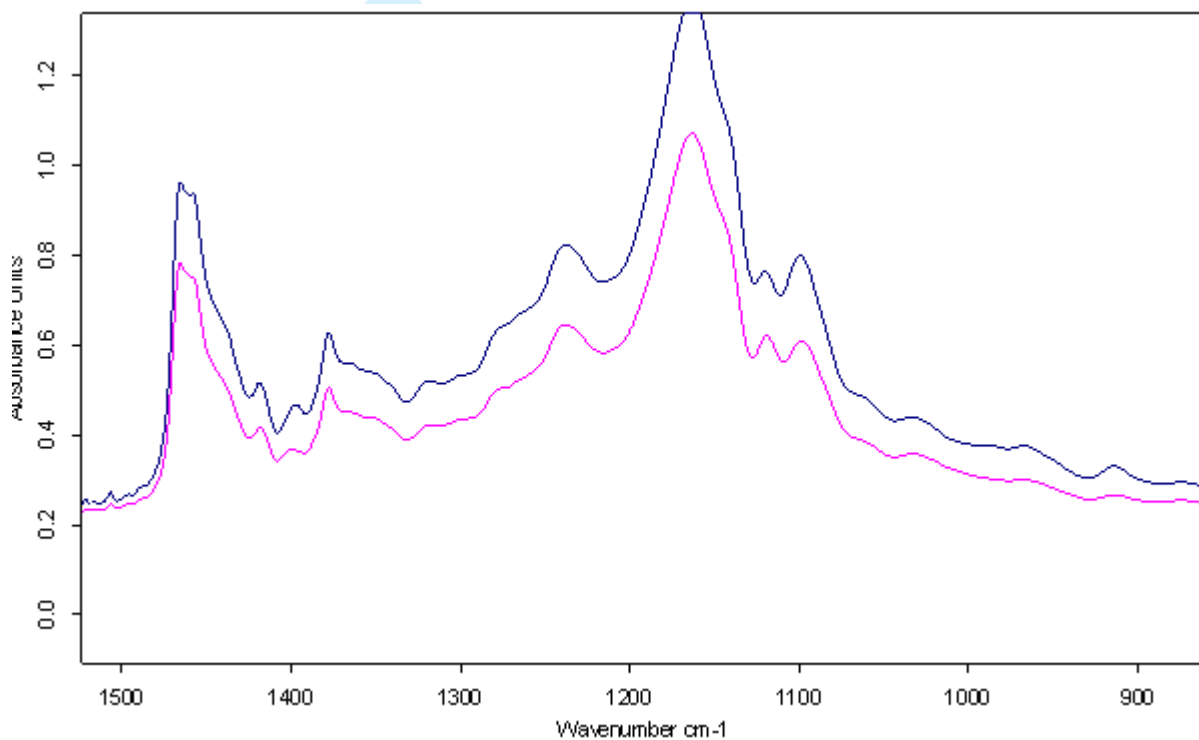
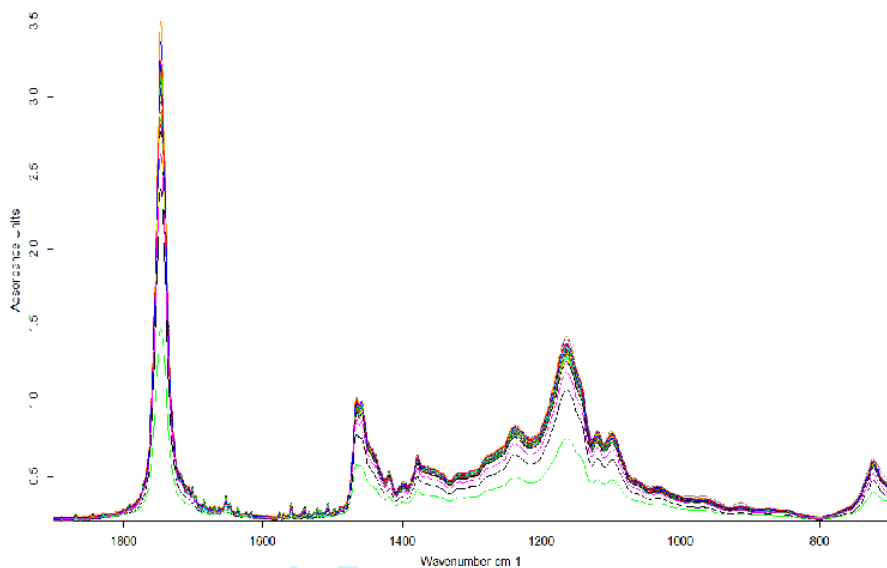


Figure 3-spectra of argan soya oils samples

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
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

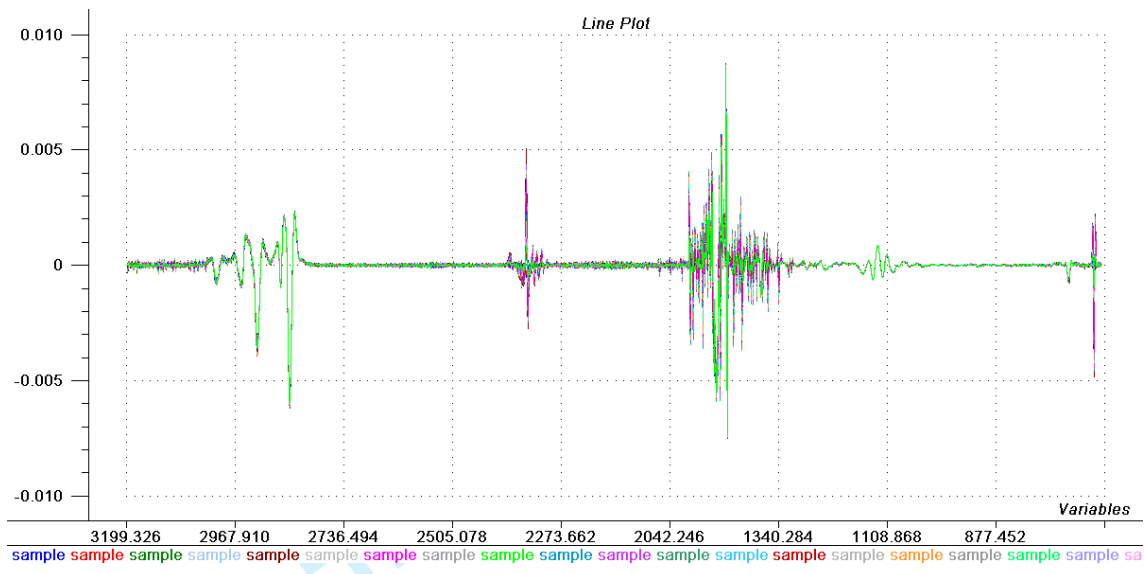


Figure 4 – spectra of argan-sunflower oil samples with *Savitzky-Golay first derivative* in 3200-600 cm^{-1}

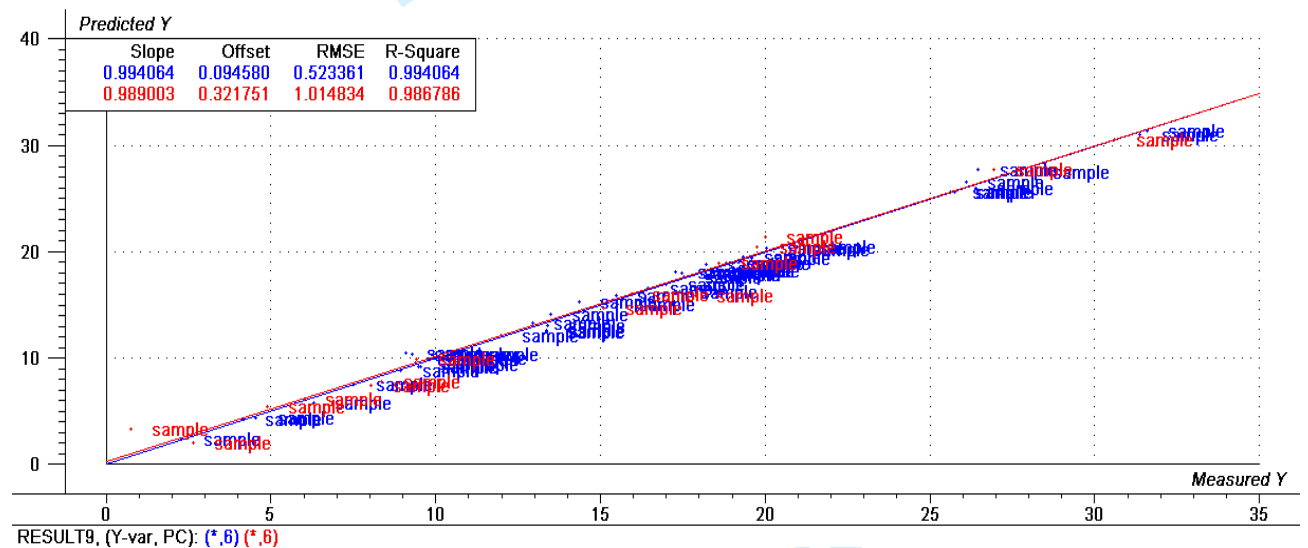


Figure 3a – Actual versus predicted sunflower oil concentration in % (w/w) for cross validation with the first savitzky- golay derivatives pretreatment

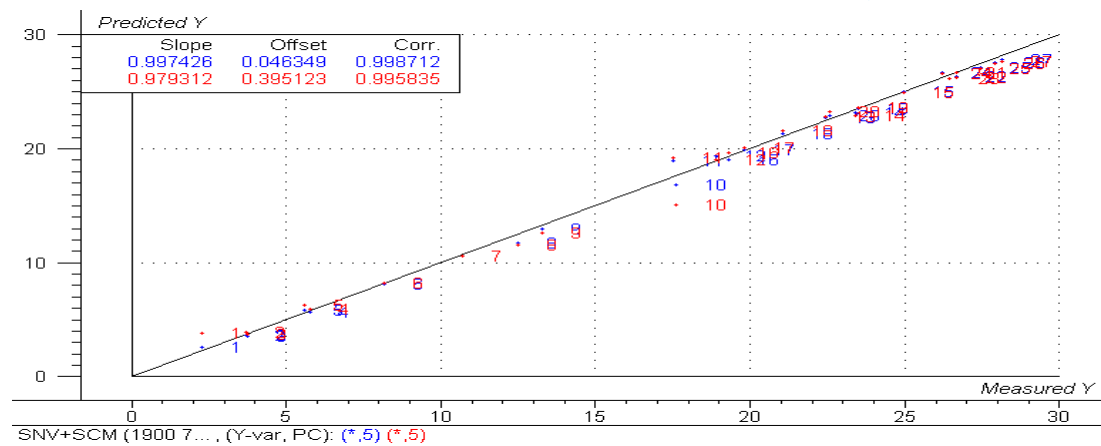


Figure 3a: Actual versus predicted soya oil concentration in % (w/w) for cross validation and calibration with the SNV+SCM pretreatment

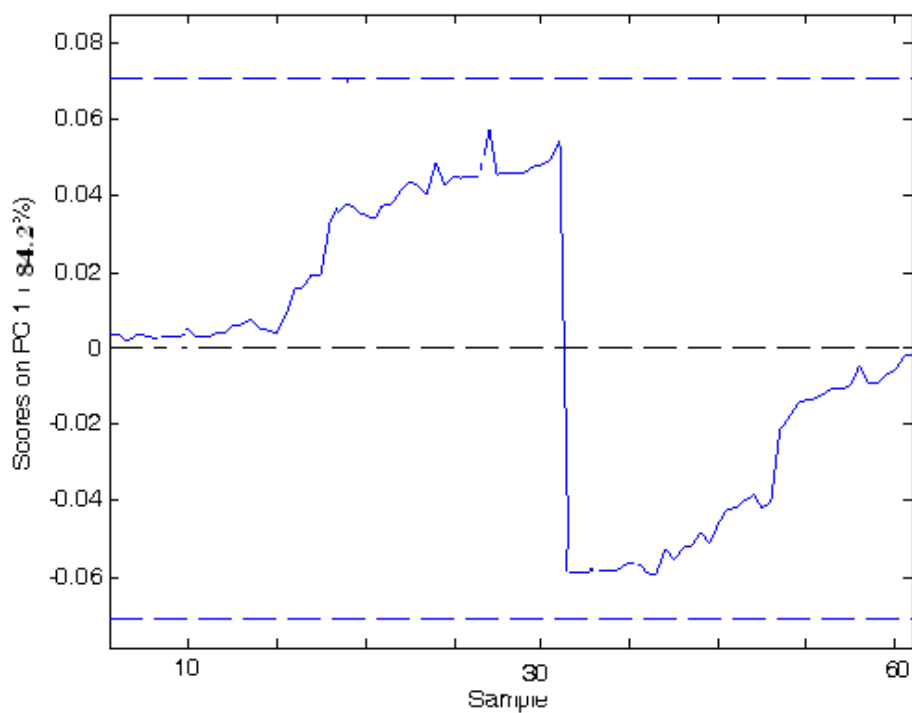


Figure 4- : Classification of oil argan adulterate with soya and sunflower oil (Cross validation: leave one out)

Table 1– Parameters for calibration (cross validation) in PLS regression models for sunflower

| Range | Pretreatment | Nombre of factors | RMSEC | RMSEP |
|-----------|--------------------------|-------------------|-------|-------|
| 4000-600 | - | 6 | 0.995 | 1.557 |
| 3000-1000 | - | 5 | 2.236 | 2.762 |
| 3000-1000 | center | 8 | 0.838 | 1.174 |
| 3000-1000 | scaled | 8 | 0.780 | 1.344 |
| 3000-1000 | 1derivatives (sgolay3-0) | 6 | 0.523 | 1.014 |
| 3000-1000 | SNV | 7 | 0.961 | 1.342 |
| 3000-1000 | MSC | 7 | 0.963 | 1.329 |
| 3000-1000 | 2derivative s golay | 4 | 1.061 | 1.408 |
| 3000-1000 | Smoothing sgolay (3-0) | 8 | 0.857 | 1.180 |
| 3000-1000 | Smoothing sgolay (3-1) | 8 | 0.857 | 1.180 |
| 3000-100 | Smoothing S.golay(5-0) | 8 | 0.874 | 1.184 |

Table 2– Parameters for calibration (cross validation) in PLS regression models for soya

| Spectral range (cm ⁻¹) | Pretreatment | Number of factor | RMSECV (%) | RMSEP (%) |
|------------------------------------|-------------------------|------------------|------------|-----------|
| 1900-700 | SNV + Smoothing average | 5 | 0.434 | 0.235 |
| 1900-700 | SNV + SMC | 5 | 0.434 | 0.235 |
| 1900-700 | SNV+ Baseline | 5 | 0.45 | 0.22 |

Actual concentration

Predicted concentration

Error

| | | |
|-------|--------|--------|
| | | |
| 4,88 | 5,665 | 0,785 |
| 6,01 | 6,235 | 0,225 |
| 2,63 | 3,031 | 0,401 |
| 8,05 | 7,544 | -0,506 |
| 8,36 | 7,454 | -0,906 |
| 9,42 | 10,133 | 0,713 |
| 15,11 | 14,728 | -0,382 |
| 15,93 | 16,399 | 0,469 |
| 17,88 | 15,861 | -2,019 |
| 18,61 | 18,87 | 0,26 |
| 19,75 | 20,278 | 0,528 |
| 20 | 21,246 | 1,246 |
| 26,94 | 27,628 | 0,688 |
| 30,6 | 30,628 | 0,028 |

Table 3 – Results for the validation set for 5 PLS component for sunflower

Table 4: parameters for prediction in PLS regression for sunflower

| Model | Elements | R-Square | Correlation | SEP | RMSEP | Slope | Offset | Bias |
|-------|----------|----------|-------------|-------|-------|-------|--------|-------|
| 5 PC | 14 | 0.990 | 0.995 | 0.834 | 0.811 | 1.004 | 0.042 | 0.109 |

Table 5-results from soya validation set

| Actual concentration | Predicted concentration | Error |
|----------------------|-------------------------|--------|
| 1.42 | 1.651 | 0.231 |
| 10.69 | 9.406 | -1.284 |
| 16.03 | 15.902 | -0.128 |
| 18.89 | 18.667 | 0.223 |
| 19.35 | 18.79 | -0.56 |
| 23.5 | 22.41 | -1.09 |
| 28.03 | 27.507 | -0.523 |
| 27.7 | 27.273 | -0.427 |

Table -6: Percent Variance Captured by PCA Model

| Principal component Number | Eigen value of Cov(X) | % Variance Captured this PC | % Variance Captured Total |
|----------------------------|-----------------------|-----------------------------|---------------------------|
| 1 | $1.27 e^{-003}$ | 83.10 | 84.20 |
| 2 | $1.27 e^{-004}$ | 7.17 | 91.37 |
| 3 | $7.83 e^{-005}$ | 6.11 | 97.48 |

For Peer Review

Classification of geographical origin of argan kernels by FTIR spectroscopy, cluster analysis and partial least squares-discriminant analysis

Fatiha EL ABADI¹, Anna de Juan², Fouzia Kzaiber¹, Romà Tauler³ and Abdelkhalek Oussama¹.

¹*Laboratoire de Spectrochimie Appliquée et Environnement, Faculté des Sciences et Techniques, Béni Mellal, Université Sultan Moulay Slimane, Maroc, fatihateto@yahoo.fr, oussamaabdelkhalek@yahoo.fr*

²*Department of analytical chemistry, Universitat de Barcelona, Diagonal, 645, 08028 Barcelona, Spain*

³*Department of environmental chemistry, IDAEA-CSIC, Jordi Girona 18-26, 08034 Barcelona, Spain.*

Abstract

The main objective of this study is the use of Fourier transform mid-infrared (FTIR) spectroscopy for discrimination of samples of argan kernels from three Moroccan regions. The content of argan oil extracted from the kernel depends on the plant genotype and on the environmental conditions. The origin of argan kernel samples could be distinguished combining the IR spectra measurement and chemometric tools for discriminant analysis. A multivariate analysis procedure based on the combined use of Hierarchical Cluster Analysis (HCA) and Partial Least Squares-Discriminant Analysis (PLS-DA) was tested for this purpose. HCA showed three distinct clusters related to the Moroccan regions and suggested the usefulness of IR spectroscopy to distinguish among kernel origins. Afterwards, PLS-DA was used for the discrimination and classification of the origin of the various argan kernels and good classification models and validation results were obtained.

The results demonstrated that the combined use of FTIR and chemometric analysis (cluster analysis and discrimination by PLS-DA) can be used to rapidly and simply determine the origine of argan kernels, a relevant task for the agroalimentary Moroccan industry.

Keywords: FTIR; Discriminant analysis; argan oil; food authentication; cluster analysis; PLS-DA.

Introduction

Argan oil is made from the Argan tree (*Argania spinosa*), endemic from southwestern Morocco protected by UNESCO. Argan oil is well known for its cosmetic, pharmaceutical and nutritional applications and has particularly unique organoleptic properties, related to its rich aroma and nutty flavour. Argan oil is considered a luxury product and is exported from Morocco around the world (Cherki et al 2006).

Authentication of food covers many different aspects, such as adulteration, characterisation, mislabelling and misleading origin. Food industry and regulatory bodies are compelled to put an increasing effort to authenticate the origin of foodstuff and to control the food constituents to ensure the sanitary safety of the product and to respond to the quality demands of the consumer (Marcos-Lorenzo et al 2002; Sayazo et al 2004), (Reid et al 2006). The origin of argan oil is a relevant information that should be authenticated, since the climatology and conditions of the different regions provide particular properties to this kind of product, as it happens with olive oils. To decrease costs and risks, an important issue could be controlling the origin of the argan kernel before the oil production step to prevent non-authenticity problems before manufacturing the commercial argan oil.

Analytical methods used to characterize argan oil components often rely on chromatographic analysis. These separation techniques (gas chromatography, GC; with mass spectrometry, GC-MS; and high-performance liquid chromatography, HPLC) have been extensively used to quantify fatty acids (Aparicio et al 2000), as well as minor compound classes such as sterols, (Alberghina, 1991) phenolic compounds (Garcia et al 2002), and hydrocarbons (Bortolomeazzi, et al 2001) in oil samples. However, routine chromatographic analysis is often time consuming and expensive and does not give straightforward information on the global quality of the sample (Hilali et al 2005). Therefore, alternative measurements cheaper, faster and non-destructive that can provide a quick answer on the origin of the sample are worth to be explored.

In this sense, mid-IR (MIR) spectroscopy has many applications because the fingerprints of many functional groups display a narrow, specific and intense signal in the MIR region (4000–400 cm^{-1}). Fourier transform infrared (FTIR) spectroscopy with attenuated total reflectance (ATR) or transmission cell accessories has been used to authenticate, identify and classify fats and oils (Lai et al., 1994). Therefore, it seems to be a potential good technique for argan oil authentication. However, although studies on oil samples have been carried out, to our knowledge, there have been no attempts to use this technique to prevent future problems

of authentication of argan oil by previous analysis of the argan kernel used to obtain this product.

Mid or near IR spectroscopy allows for fast monitoring of oil samples without the need of sample pretreatment. However, the use of this analytical measurement should go accompanied by the use of multivariate data analysis tools to provide good models of calibration or classification of samples that can be later used to predict oil component contents or origins of new unknown samples, respectively. Exploring the existence of natural clustering patterns related to the origin of foodstuff samples can be made by unsupervised methods, which identify the group objects on the basis of similarities among the samples (Massart et al 1983),, [Forina et al 2005](#)). Common methods for unsupervised pattern recognition are Principal Component Analysis (PCA) and cluster analysis (CA), which can give a first insight about the possibility of using a certain measurement, e.g., FTIR spectra, to detect groups of samples according to a certain property; in this case, the origin of the argan kernel. If these groups are detected, classification tools, which build models that establish the relationship between the analytical measurement and the sample class (e.g., based on the origin) can be built, validated and used to predict the origin of new unknown samples. Diverse chemometric methods for classification have been successfully applied to the authentication of olive fruit and olive oil origin (Dahlberg, 1997), Bendini et al 2007), [M. de Luca, et al \(2011\)](#) , [Terouzi et al 2011](#)). [To comment a few, Marini et al. \(2004\)](#) authenticated e.v. olive oil varieties using linear discriminant analysis (LDA) and back propagation artificial neural networks (BP-ANN), [or Downey et al \(2003\)](#) studied the visible and NIR spectra of e.v. olive oils and correctly classified 94% of the samples according to the geographical origin, using PLS-LDA and *k*-nearest neighbour method KNN.

The aim of the present work is to develop a new application of the FTIR technique combined with unsupervised and supervised classification methods as a rapid, inexpensive and nondestructive tool, useful to determine the geographic origin of argan kernels, later on used for argan oil manufacturing. Verifying the declared origin or determining the origin of unidentified argan kernels is a relevant problem because of the increasing use and export of this natural product and the benefit that can report detecting fraud in authentication before the argan oil product is manufactured. The present work will test the combined use of FTIR spectroscopy with multivariate data analysis methods, such as Hierarchical Cluster Analysis (HCA) and PLS-DA for the authentication of the origin of argan kernels samples from three Moroccan regions: Agadir, Guelmim and Ait Baha.

2. Materials and methods

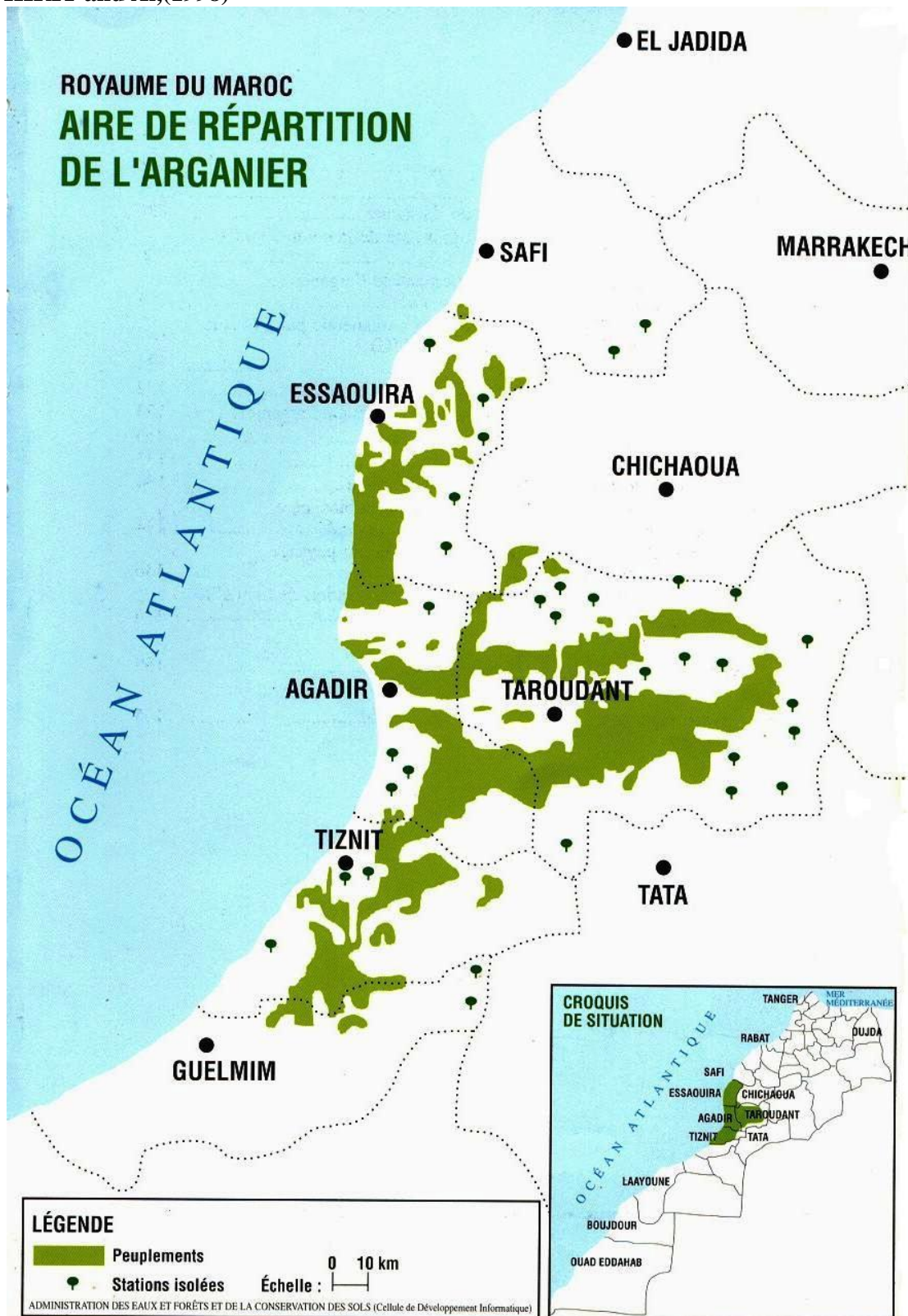
2.1. Samples

30 argan kernel samples were obtained from three Moroccan regions during the 2010 harvest season (10 samples per each region). The collection method (pulping, crushing of nut and obtaining the argan kernels), the storage conditions and preservation were the same for all the samples used. Natural argan kernel samples were analyzed, without being subject to roasting treatments. Investigation was focused on argan samples picked in the zones named Ait Baha (AITB), Agadir (Agt) and Guelmim (Glm). (Fig. 1) Geographic characteristics such as altitude, latitude and longitude as well as the mean rainfall of these ecoregions are summarized in Table 1.

Table 1: Climate and geographical properties of the Moroccan regions of argan production.

| Ecoregions | Latitude N | Longitude W | Altitude (m) | Rainfall (mm) |
|------------|------------|-------------|--------------|---------------|
| Agadir | 30°41' | 9°33' | 150–350 | 300 |
| Guelmim | 28°58' | 10°3' | 286 | 217 |
| Ait baha | 29°33' | 8°58' | 340-1380 | 218 |

Figure 1: Chart of the surface of distribution of the argan in Moroccan South-west (Me HIRIT and Al,(1998)



2.2. FT-IR analysis

The MIR measurements were acquired with a Bruker Vector 22 FTIR spectrometer. This instrument was equipped with an Attenuated Total Reflectance (ATR) sampling accessory (Diamant crystal) and a deuterated triglycine sulphate (DTGS) detector. The ATR crystal was carefully cleaned with pure ethanol to eliminate the presence of oil/fat residues between measurements to ensure a clean crystal surface to obtain the best possible sample spectra. Kernel samples were analysed and an FTIR spectrum per sample was recorded in the wavenumber range from 4000 to 600 cm^{-1} . A total sum of 104 scans was collected and averaged per each sample. The spectral resolution was 1cm^{-1} . The OPUS (5.5) software program was employed for spectral data collection.

2.3.Data Analysis:

Two data subsets were prepared with the 30 FTIR spectra collected from the argan samples, a calibration set with 21 samples and a validation set with 9. The Kennard-Stone method was used for data subset construction (Kennard, 1969). Both calibration and validation sets were balanced, i.e., the same number of samples per class was selected. For the different data analysis performed, different pretreatments were tested on the FTIR spectra collected, such as Standard Normal Variate (SNV) and 1st and 2nd derivative by Savitzky-Golay method, with different options of data window and polynomial order (Azzouz (2003), Candolfi (1999)).

Hierarchical cluster analysis (HCA) was first applied to the calibration set to detect the possibility to distinguish natural groups of argan kernel samples linked to the different Moroccan regions from the FTIR spectra measured. In HCA, each sample is initially assumed to be a lone cluster. Then, samples are joined sequentially to form clusters according to the distance among them. First, the two closest samples are joined to form a cluster and in subsequent hierarchical steps all the rest are clustered until a single group formed by all samples analyzed is formed. The final results are represented as a hierarchical tree structure with all the clustering steps represented, the so-called dendrogram (Jolliffe et al 1989, Massart (1983)). In this work, the average linkage hierarchic agglomerative cluster algorithm with Euclidean distance was adopted to perform the analysis of the FTIR data.

Once seen the possibility to use FTIR spectra as a way to distinguish among argan kernel samples from different regions, Partial Least Squares-Discriminant Analysis (PLS-DA) was applied to build a classification model that could help to assign a geographic origin to unknown argan kernel samples. PLS-DA is a classification method based on the multivariate calibration PLS method, which looks for the model that relates \mathbf{X} (matrix with predictor

information) and **Y** (matrix with properties to be predicted) by using a set of latent variables, calculated to represent the directions of maximum covariance between **X** and **Y**. In the PLS-DA method, the **X** matrix contains the instrumental responses of the samples, e.g. the FTIR spectra, and the **Y** matrix the information on the sample class membership, e.g., origin of the argan kernel. In PLS-DA, samples belonging to a class are coded with 1 and samples out of the class with 0. PLS-DA aims at finding the relationship between **X** and **Y**, i.e., the relationship between the instrumental responses and the class membership of the samples ([Barker \(2003\)](#), [de Luca, \(2011\)](#)).

In this study, PLS-DA was applied using the PLS2 algorithm, which allows setting a **Y** matrix with several variables, one per each of the classes represented, Ait Baha, Guelmim and Agadir. For each class, the **y** variable is set equal to 1 for the samples belonging to the class and to 0 for those not belonging to the class.

The size of the PLS-DA model was chosen by full cross-validation. The number of latent variables was chosen as the best compromise between a low root mean square error of prediction on the cross-validated samples (RMSECV), and a low number of misclassified samples, always trying to keep the smallest possible model. The model was finally validated on an external set of nine validation samples.

The UNSCRAMBLER 10 package (CAMO, Norway) was used for all data preprocessing and analysis.

3.Results and discussion

FTIR spectra of 30 argan kernel samples, produced in different locations of Morocco were recorded and divided into a calibration set of 21 samples and a prediction set of 9 samples, as mentioned in the previous section. Fig. 2 shows the mean spectra of the three classes studied. The differences among the FTIR spectra shown were small and occurred only in limited regions of the wavenumber range .The band of water around 2994 cm^{-1} has not been removed for this study. Peaks around $2800\text{--}3080\text{ cm}^{-1}$ are due to hydrogen stretching mode, the large peak between $1700\text{ and }1800\text{ cm}^{-1}$ could be attributed to C=O stretching and C–O–C stretching and C–H bending in the region of $900\text{--}1400\text{ cm}^{-1}$ could be easily observed ([Tay et al., 2002](#)).

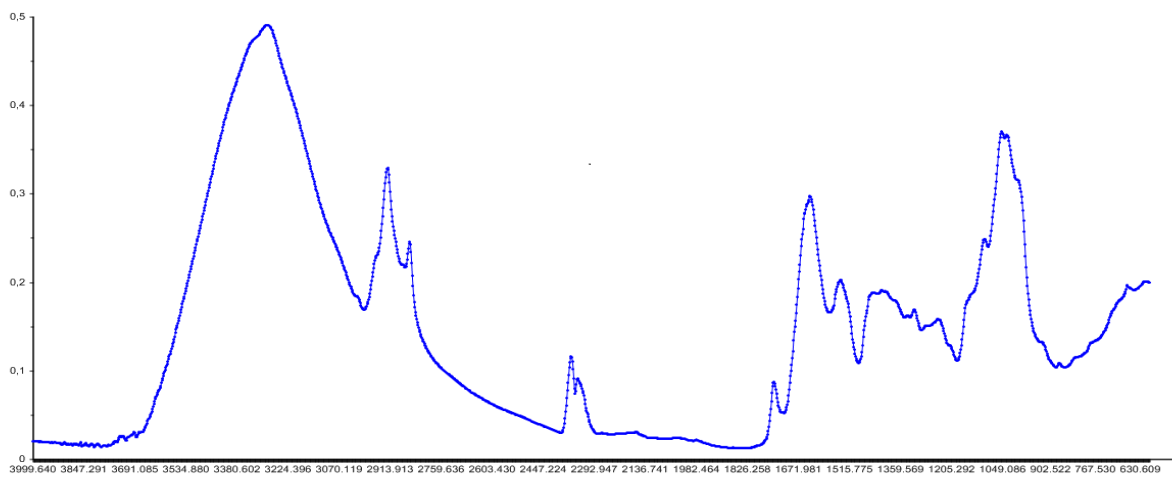
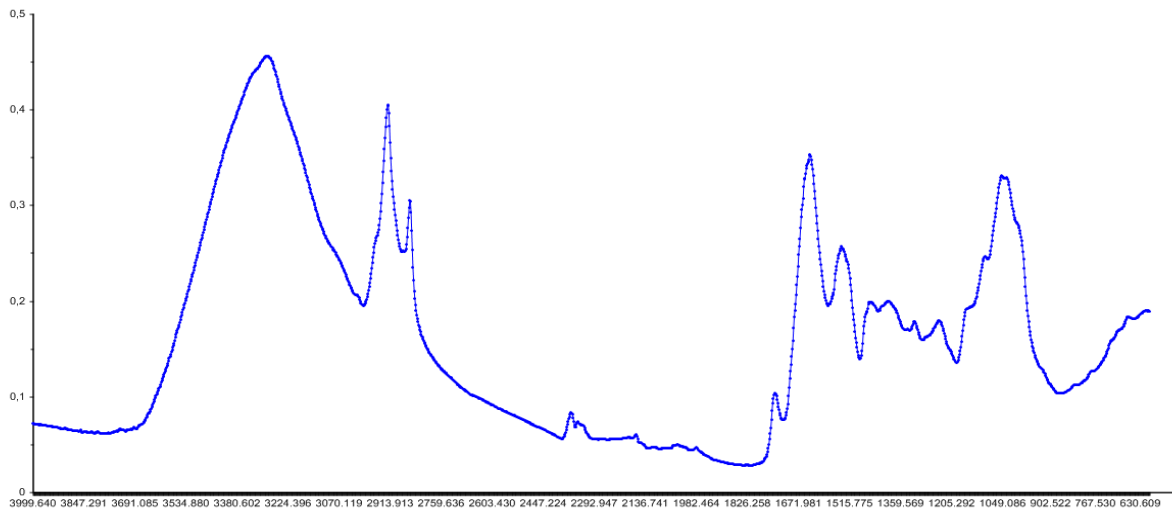
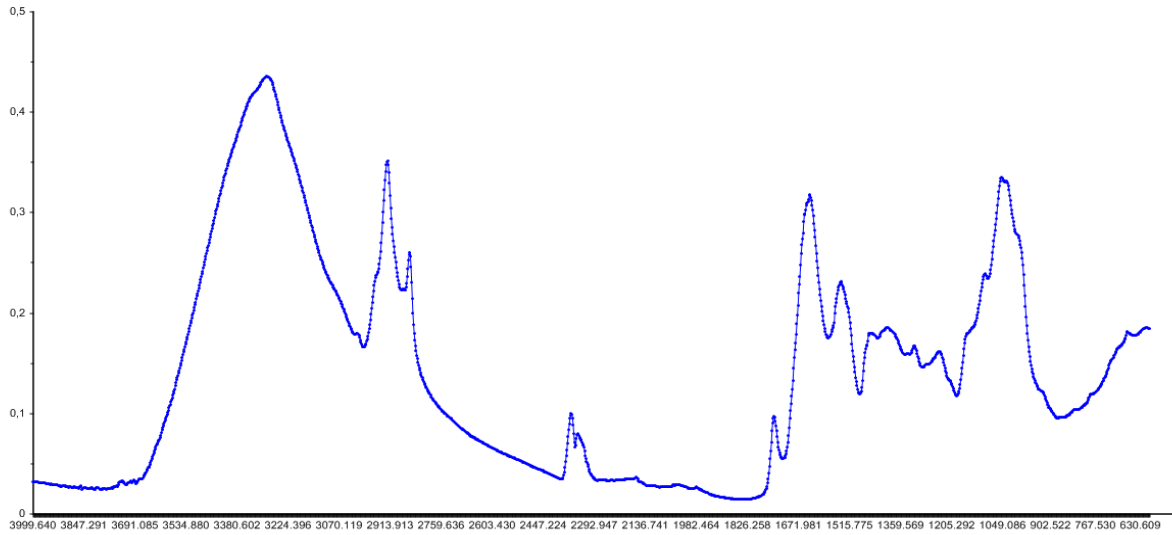


Fig. 2. Mean FTIR spectrum of Agadir, Ait baha and Guelmim respectively

With the aim to obtain more clear information from the FTIR spectral data, the spectra were first subject to mathematical pretreatment. In particular, Standard Normal Variate (SNV) and Derivative pretreatments were tested for HCA and PLS-DA purposes (Iñón et al. 2003; Maggio et al., 2009) (Azzouz (2003), Candolfi (1999)).

3.1. Hierarchical Cluster Analysis (HCA)

HCA is an unsupervised technique that uses the information obtained from measured variables to reveal natural clustering structures existing between the studied samples (Massart, 1983)

The results presented for cluster analysis have been obtained on first derivative FTIR spectra obtained applying the Savitzky-Golay method with a 2nd order polynomial and 5-point window smoothing (Savitzky,1964). Figure 3 shows the resulting dendrogram obtained from HCA using the average linkage hierarchic agglomerative cluster algorithm with Euclidean distance on the derivative FTIR spectra.

Three clusters can be clearly identified. The first group contains the samples collected in the territory of Ait baha, the second the samples from Guelmim, an area characterised by hilly uplands and high temperature range, and the last cluster consists of the samples picked in the territory of Agadir. The dendrogram shows that FTIR spectra allow for a very clear definition of the origin of the three types of argan kernels. As can be seen in the dendrogram, the three origins lead to very compact clusters, formed always by samples from the same origin. These three clusters are clearly separated from one another.

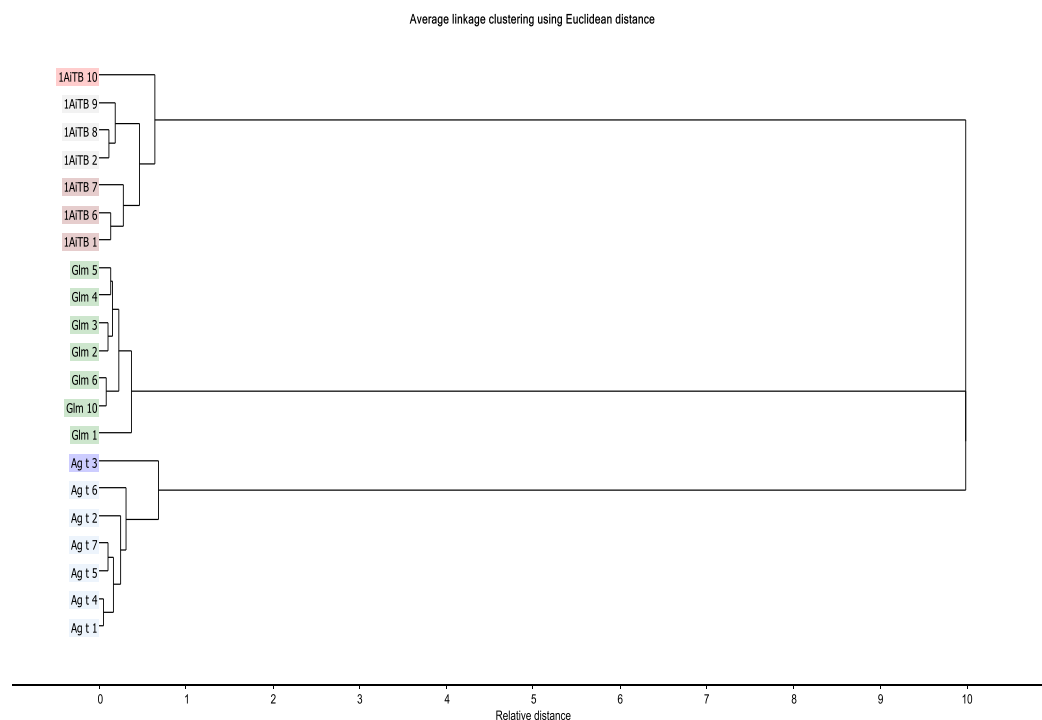


Fig. 3: HCA dendrogram of Ait baha (1AITB), Agadir (Agt), and Guelmim (Glm) argan kernels using the average linkage algorithm and computing Euclidean distances.

The clear natural clustering pattern of the argan kernels from the three Moroccan origins seen in the dendrogram suggests the possibility to build classification models based on FTIR information from these kernels to predict the origin of unknown argan samples.

3.2. Classification model

For classification purposes, PLS-DA was applied to the set of 21 calibration samples with the aim of building a classification model useful for the later assignment of new argan kernels samples of unknown origin. A PLS2 algorithm was used with the \mathbf{X} matrix of FTIR spectra and a \mathbf{Y} matrix containing three variables, one per each class (argan kernel origin). Each class was defined with a binary \mathbf{y} -variable, i.e., setting 1 for samples in the class and 0 for samples out of the class. Different data pretreatments along the row (spectra) direction were tested, such as Standard Normal Variate (SNV) and 1st and 2nd derivative with different polynomial orders and point window sizes. For these treatments, additional variable centering and autoscaling was also tested. The choice of the best classification model built on the calibration samples was based on the results obtained by full cross-validation, which are summarised in Table 2. For each of the models built, global parameters such as the number of

latent variables in the model and the Y explained variance are provided. For each of the three classes in the model: Ait Baha, (-aitb), Agadir (-agt) and Guelmim (-glm), figures such as the RMSEP (root mean square error of prediction) and the squared correlation coefficient between y_{pred} and y_{measured} ($R^2\text{CV}$) are given. In addition, the classification quality for each class within each model is evaluated by providing the number of misclassifications, separating false positives (samples identified as in the class, when they are not) and false negatives (samples identified as out of the class when they are in). For this purpose, a threshold value of 0.5 has been chosen and samples with a predicted y value above 0.5 are considered in the class and below 0.5 as out of the class.

Table 2. Statistical parameters obtained from full cross-validation on the PLS-DA model.

| Pretreatment ^(*) | Class | Number of factors | Y expl. var. (%) | RMSECV | $R^2\text{CV}$ | False + | False - |
|-----------------------------|-------|-------------------|------------------|--------|----------------|---------|---------|
| Raw data (none) | -aitb | 3 | 47.06 | 0.2524 | 0.756 | 2 | 0 |
| | -agt | | | 0.4875 | 0.605 | 2 | 3 |
| | -glm | | | 0.3236 | 0.7 | 1 | 1 |
| Raw data (CN) | -aitb | 6 | 62.27 | 0.2266 | 0.786 | 0 | 0 |
| | -agt | | | 0.3939 | 0.9 | 1 | 4 |
| | -glm | | | 0.2360 | 0.294 | 1 | 1 |
| SNV C | -aitb | 4 | 54.42 | 0.2457 | 0.746 | 1 | 1 |
| | -agt | | | 0.4733 | 0.7 | 4 | 2 |
| | -glm | | | 0.3011 | 0.164 | 2 | 1 |
| SNV CN | -aitb | 6 | 63.54 | 0.2412 | 0.767 | 0 | 1 |
| | -agt | | | 0.4098 | 0.2 | 3 | 2 |
| | -glm | | | 0.2366 | 0.346 | 0 | 0 |
| 2D2O5S C | -aitb | 7 | 77.26 | 0.1879 | 0.869 | 0 | 0 |
| | | | | | 0.785 | | |
| | | | | | 0.7 | | |

| | | | | | | | |
|-----------|-------|---|-------|--------|-------|---|---|
| | -agt | | | 0.3349 | 0.642 | 2 | 1 |
| | -glm | | | 0.2366 | 2 | 0 | 1 |
| | | | | | 0.815 | | |
| | | | | | 6 | | |
| 2D2O5S CN | -aitb | 5 | 66.96 | 0.2431 | 0.756 | 0 | 1 |
| | -agt | | | 0.3545 | 5 | 2 | 1 |
| | -glm | | | 0.2235 | 0.468 | 0 | 1 |
| | | | | | 0 | | |
| | | | | | 0.782 | | |
| | | | | | 6 | | |
| 2D2O7S C | -aitb | 7 | 76.46 | 0.1722 | 0.891 | 0 | 0 |
| | -agt | | | 0.3079 | 0 | 1 | 0 |
| | -glm | | | 0.2335 | 0.634 | 0 | 1 |
| | | | | | 8 | | |
| | | | | | 0.771 | | |
| | | | | | 1 | | |
| 2D2O7S CN | -aitb | 5 | 67.26 | 0.2024 | 0.828 | 0 | 0 |
| | -agt | | | 0.3702 | 5 | 2 | 2 |
| | -glm | | | 0.2679 | 0.470 | 0 | 1 |
| | | | | | 9 | | |
| | | | | | 0.722 | | |
| | | | | | 6 | | |
| 2D2O9S C | -aitb | 7 | 75.47 | 0.1660 | 0.885 | 0 | 0 |
| | -agt | | | 0.2990 | 0 | 1 | 0 |
| | -glm | | | 0.2367 | 0.635 | 0 | 1 |
| | | | | | 1 | | |
| | | | | | 0.752 | | |
| | | | | | 7 | | |
| 2D2O9S CN | -aitb | 6 | 69.29 | 0.1820 | 0.850 | 0 | 0 |
| | -agt | | | 0.3493 | 4 | 2 | 2 |
| | -glm | | | 0.2411 | 0.481 | 0 | 0 |
| | | | | | 8 | | |
| | | | | | 0.748 | | |
| | | | | | 5 | | |
| 1D2O5S C | -aitb | 7 | 64.39 | 0.2110 | 0.812 | 0 | 0 |
| | -agt | | | 0.3807 | 9 | 1 | 3 |
| | -glm | | | 0.2694 | 0.433 | 0 | 1 |
| | | | | | 3 | | |
| | | | | | 0.697 | | |
| | | | | | 7 | | |
| 1D2O5S CN | -aitb | 6 | 80.73 | 0.1675 | 0.883 | 0 | 0 |

| | | | | |
|------|--------|-------|---|---|
| -agt | 0.2850 | 6 | 1 | 0 |
| -glm | 0.1881 | 0.672 | 0 | 0 |
| | | 7 | | |
| | | 0.862 | | |
| | | 6 | | |

(*) C and CN: column centered and column autoscaled data, respectively. For derivative pretreatment, e.g., 1D2O5S, 1D or 2D mean 1st or 2nd derivative, respectively, 2O means 2nd order polynomial and 5S, five points of window size.

Irrespective from the pretreatment used, the Agadir class is always the most difficult to model, which can be clearly seen through the RMSEP, r^2 value and the number of misclassifications (false positives and false negatives). The other two classes show satisfactory classification rates with almost all pretreatments tested. Trying to look for the best global results for all classes, the derivative pretreatment seems to be more appropriate than SNV or to the work with raw data, according to the values of RMSEP, R^2CV and number of misclassifications. The pretreatments along the rows, i.e., SNV or derivatives, seem to influence more the quality of the final results than column centering or autoscaling. Within the different derivative options tested, Table 2 shows that the first derivative using second order polynomial, five point window and autoscaling provides the best results in terms of number of misclassifications, RMSEP and R^2CV and has a high Y variance explained. Other options, such as second derivative with second order polynomial and 7 or 9 points window with column centering are also appropriate alternatives in terms of classification rate but seem to be slightly worse when RMSEP and R^2CV are observed.

Focusing on the PLS-DA model finally selected, figure 4 shows the PLS-DA score plot for the calibration samples. Using two latent variables, 72% of Y variance is already explained. Three regions (group) can be clearly identified, each of them representing one type of argan kernels region. This result clearly shows the potential ability of this technique to discriminate among the three argan kernel origins used in this study. At the same time, it is worth noticing that the Agadir class is placed between the other two. This explains the higher difficulty in distinguishing these samples from the rest of the classes.

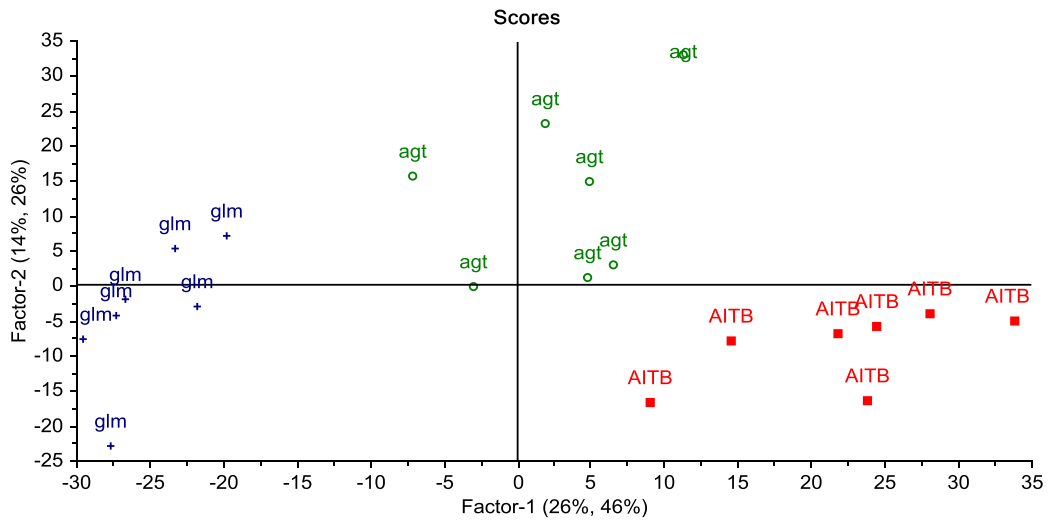
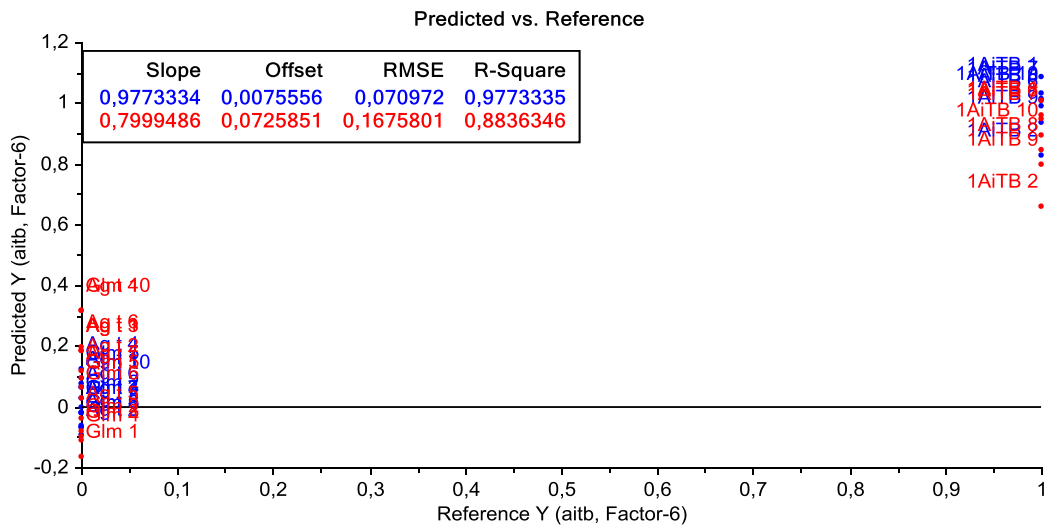


Fig. 4: Score plot of PLS-DA model for argan kernels of Ait baha (AITB), Agadir (agt) and Guelmim (glm) from autoscaled data pretreated with first derivative with five point window.

Figure 5 shows the y predicted vs. y measured results for each one of the classes in this classification model. The red dots indicate the cross validation results for the calibration samples, which are the ones presented in Table 2.



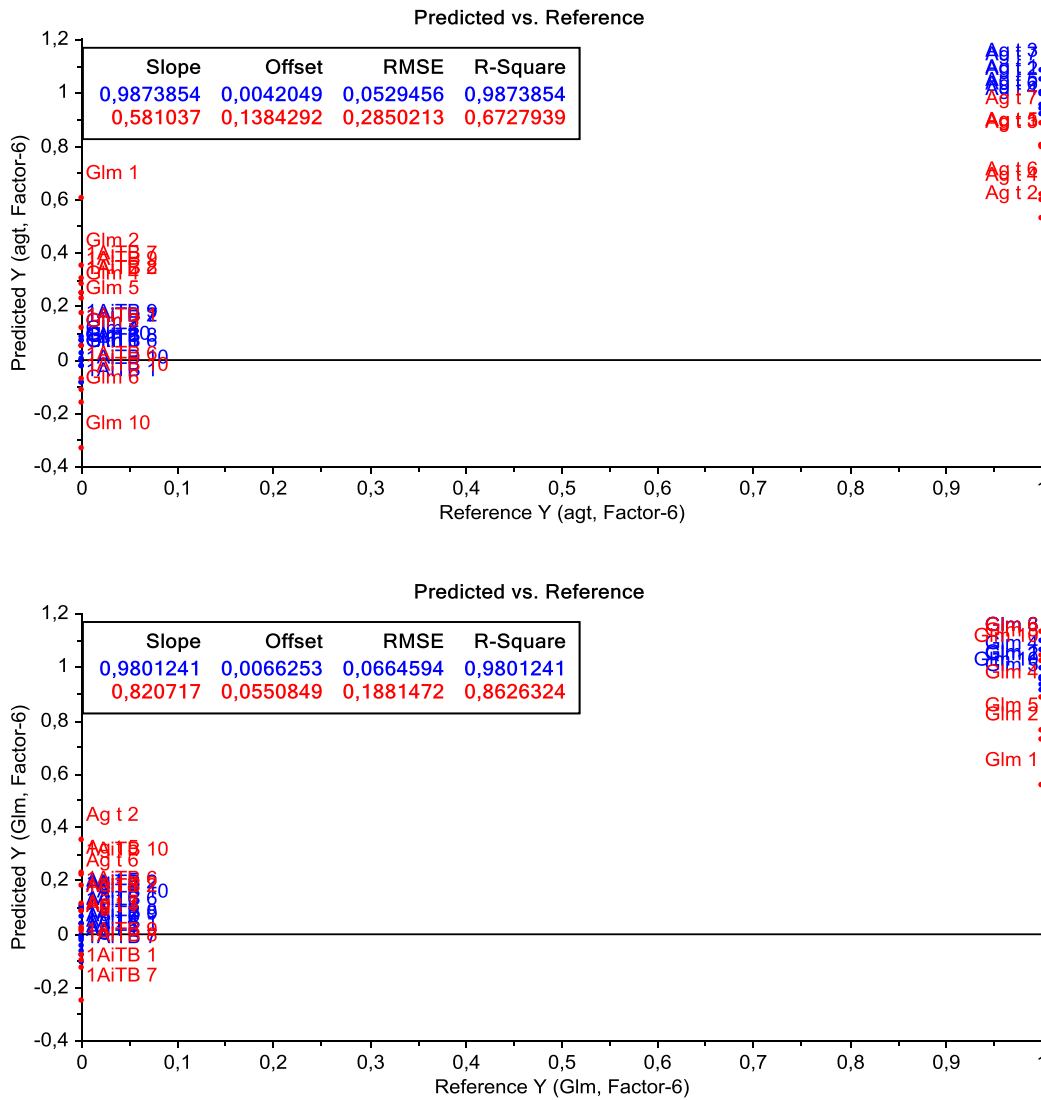


Fig. 5: Predicted value versus measured value of Ait baha , Agadir and Golmume of PLS –DA model.

The best results are found for the Ait Baha class, where calibration and cross-validation results are very similar. Guelmimclass is also well modeled in the cross-validation results although the y predicted values are more spread (less compact in ‘in and out’ of the class groups) than the calibration results. The larger worsening when comparing calibration and cross-validation results in the Agadir class, as expected, where slope, RMSE and r^2 worsen significantly from calibration to validation results. Nevertheless, the classification rate is still satisfactory, although some of the samples remain very close to the class separation limit. Within the three classes, Agadir and Guelmim are mainly responsible for missclassifications being sample 2 of Agadir close to be a false – in its own class variable and a false - in Guelmim class variable; the opposite situation occurs with sample 1 from Guelmim.

Classification of new samples

The chosen classification model was applied on the FTIR external validation set, formed by nine new argan kernels samples, three from each class. This step allowed testing the ability of the built model to classify new samples into the classes previously established. Table 3 shows the classification results with the comparison between the predicted values and the expected theoretical value for each sample in each class variable.

Table 3: Classification results of argan kernel validation samples. The bold values show the samples that confirm the expected class membership.

| | AIT BAHA | | AGADIR | | GUELMIM | |
|--------------|---------------|-----------|---------------|-----------|---------------|-----------|
| | Prediction | Reference | Prediction | Reference | Prediction | Reference |
| Ait b | 1,0026 | 1 | -0.2237 | 0 | 0.2211 | 0 |
| Aitb | 0,9630 | 1 | 0.0106 | 0 | 0.0262 | 0 |
| Ait b | 1,0131 | 1 | -0.0174 | 0 | 0.0043 | 0 |
| Agt | 0,0203 | 0 | 0.9604 | 1 | 0.0192 | 0 |
| Agt | 0,0969 | 0 | 0.5999 | 1 | 0.3030 | 0 |
| Agt | 0,0223 | 0 | 0.8913 | 1 | 0.0862 | 0 |
| Glm | -0,0144 | 0 | 0.0394 | 0 | 0.9750 | 1 |
| Glm | 0,0713 | 0 | -0.3094 | 0 | 1.2381 | 1 |
| Glm | 0,0556 | 0 | -0.2679 | 0 | 1.2122 | 1 |

Figure 6 a), b) and c) show in more detail the results of Table 3, with the error in prediction bar for the samples of each of the classes. Fig. 6a shows the classification results for the variable related to the Ait baha class. This is the best defined class and samples in the class are well above the threshold value of 0.5, even considering the error associated with the prediction, and samples out of the class are very well recognized too. Figure 6c shows the correct assignment of Guelmim samples into its class. Samples out of the class are classified correctly. In figure 6b, the results for the Agadir class are shown. Although the nominal y predicted values for the Agadir samples classify them in the class, sample 5 goes out of limits

when the error associated with the prediction is considered. This confirms the high difficulty linked to the proper modeling of this class. Besides, samples out of the class show a higher spread in the y values predicted as well, showing the lower compactness within the sample groups in and out of the class.

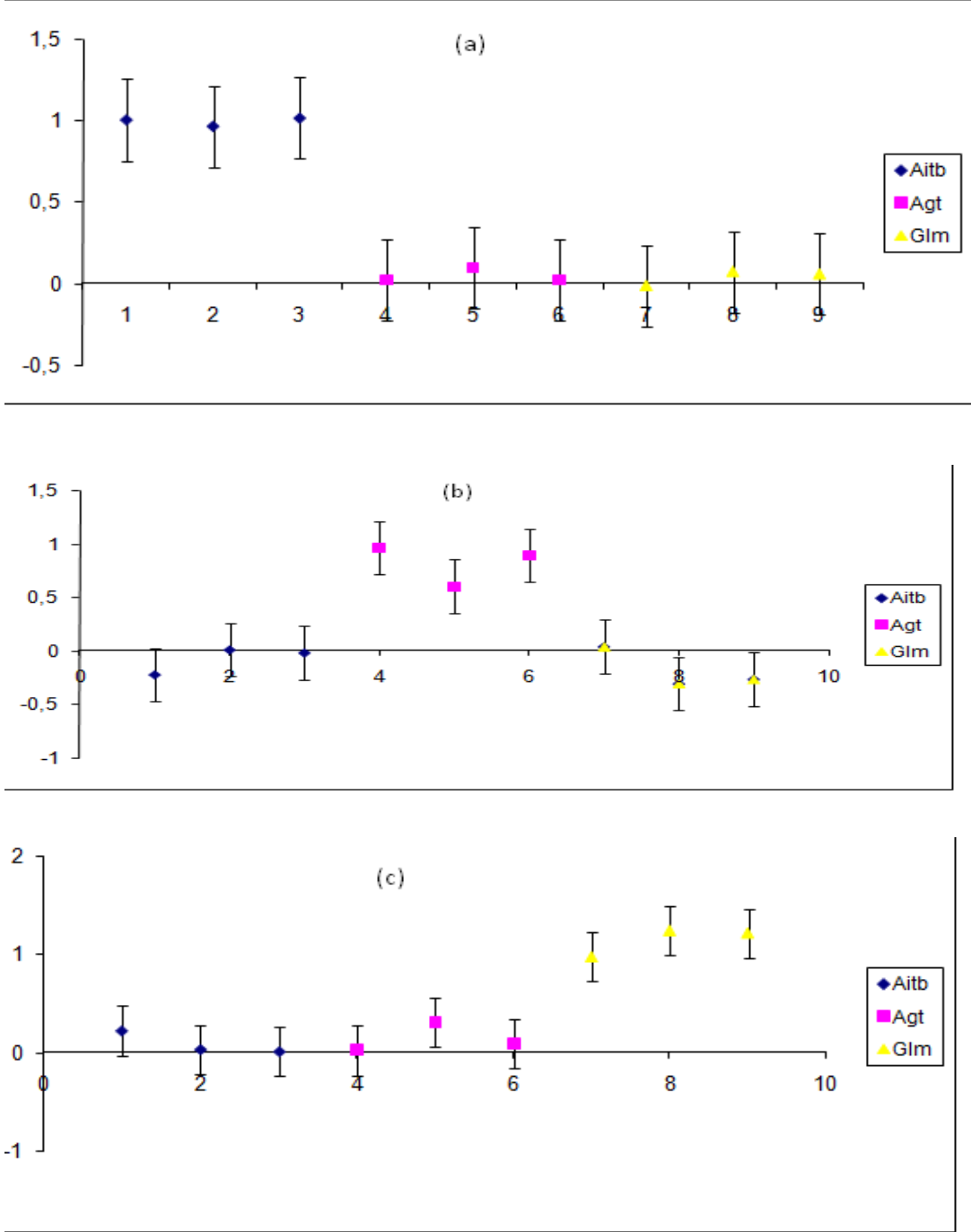


Fig. 6. Predictions for the validation set of argan pulpe: (a) class Ait baha; (b) class Agadir and (c) for class Guelmim.

Conclusions

In this paper, HCA and later PLS-DA were used in order to build a classification model to check the origin of argan kernels from FTIR measurements. The combined use of this analytical technique and adequate chemometric methods has proven to be an efficient strategy to discriminate the argan kernels according to their geographic origin without the need of further chemical analysis methods. The three regions studied could be clearly classified, although the Agadir region was more difficult to be modeled. This method can be proposed as a suitable alternative for quality control of argan kernels and determination of their geographical origin, with the special value that this authentication strategy can be performed before the manufacturing of the final argan oil.

References

- Alberghina, G.; Caruso, L.; Fisichella, S.; Musumarra, G. (1991) Geographical classification of Sicilian olive oils in terms of sterols and fatty-acids content. *J. Sci. Food Agric.*, 56, 445-455.
- Aparicio, R.; Aparicio-Ruiz, R. (2000). Authentication of vegetable oils by chromatographic techniques. *J. Chromatogr.*, 881, 93-104.
- Azzouz, T. Puigdomènech, A., Aragay, M. and Tauler, R. (2003) Comparison between different data pre-treatment methods in the analysis of forage samples using near-infrared diffuse reflectance spectroscopy and partial least-squares multivariate calibration method. *Anal. Chim. Acta*, 484 121.
- Barker, M., Rayens, W. (2003). Partial least squares for discrimination. *Journal of Chemometrics*, 17(3), 166–173.
- Bendini, A., Cerretani, L., Di Virgilio, F., Belloni, P., Bonoli-Carbognin, M., & Lercker, G. (2007). Preliminary evaluation of the application of the FTIR spectroscopy to control the geographic origin and quality of virgin olive oils. *Journal of Food Quality*, 30, 424–437.
- Bortolomeazzi, R.; Berno, P.; Pizzale, L.; Conte, L. S. 2001. Sesquiterpene, alkene, and alkane hydrocarbons in virgin olive oils of different varieties and geographical origins. *J. Agric. Food Chem.*, 49, 3278-3283.
- Candolfi, A. de Maesschalck, R. Jouan-Rimbaud, D., Hailey, P.A., Massart, D.L. The influence of data pre-processing in the pattern recognition of excipients near-infrared spectra. *J. Pharm. Biomed. Anal.* 21 (1999) 115.
- Cherki, M., Berrougui, H., Drissi, A., Adlouni, A., & Khalil, A. (2006). Argan oil: Which benefits on cardiovascular diseases? *Pharmacological Research*, 54, 1–5.
- Dahlberg, D. B., Lee, S. M., Wenger, S. J., & Vargo, J. A. (1997). Classification of vegetable oils by FT-IR. *Applied Spectroscopy*, 51, 1118–1124.
- Deluca, M., Terouzi, W., Gaetano, Oussama, A. (2011). Derivative FTIR for cluster analysis and classification of moroccan oil olive. *Food chemistry*, 124, 1113-1118.
- Downey, G., McIntyre, P., & Davies, A. N. (2003). Geographic classification of extra virgin olive oils from the eastern Mediterranean by chemometric analysis of visible and near-infrared spectroscopic data. *Applied Spectroscopy*, 57(2), 158–163.
- Forina, M., Oliveri, P., Lanteri, S., & Casale, M. (2008). Class-modeling techniques, classic and new, for old and new problems. *Chemometrics and Intelligent Laboratory Systems*, 93, 132–148.
- El Aich, A., Bourbouze, A., Chiche, J., Yessef, M., (2005) Diagnostic on the PDPEO development project (rangelands and animal husbandry in the Oriental) carried out by the students
- Garcia, A.; Brenes, M.; Romero, C.; Garcia, P.; Garrido, A. (2002). Study of phenolic compounds in virgin olive oils of the Picual variety. *Eur. Food Res. Technol.*, 215, 407-412
- Hilali, M.; Charrouf, Z.; Soulhi, A.; Hachimi, L.; Guillaume, D. (2005) Influence of Origin and Extraction Method on Argan Oil Physico-Chemical Characteristics and Composition *J. Agric. Food Chem.*, 53, 2081-2087

Iñón, F. A., Garrigues, J. M., Garrigues, S., Molina, A., & de la Guardia, M. (2003). Selection of calibration set samples in determination of olive oil acidity by partial least squares-attenuated total reflectance-Fourier transform infrared spectroscopy. *Analytica Chimica Acta*, 489, 59–75.

Jolliffe, T., O.B. Allen, B.R. Christie, 1989. Comparison of variety means using cluster analysis and dendrograms, *Experimental Agriculture*, 25: 259-269.

R.W. Kennard and L.A. Stone, Computer aided design of experiments, *Technometrics* 11 (1969) 137-148

Lai, Y. W., Kemsley, E. K., & Wilson, R. H. (1994). Potential of Fourier transform infrared spectroscopy for the authentication of vegetable oils. *Journal of Agricultural and Food Chemistry*, 42, 1154–1159.

M'hirit ,O., Benzyane ,M., Benchekroun ,F., El yousfi, M., Bendaanoun (1998). L'arganier, une espèce fruitière à usage multiples, Ouvrage collectif Edition Mardaga (Belgique) ,150 p

Maggio, R. M., Kaufman, T. S., Del Carlo, M., Cerretani, L., Bendini, A., Cichelli, A., et al. (2009). Monitoring of fatty acid composition in virgin olive oil by Fourier transformed infrared spectroscopy coupled with partial least squares. *Food Chemistry*, 114, 1549–1554.

Marcos-Lorenzo, I., Pérez-Pavón, J. L., Fernández-Laespada, M. E., García-Pinto, C., & Moreno-Cordeo, B. (2002). Detection of adulterants in olive oil by headspace mass spectrometry. *Journal of Chromatography A*, 945, 221–230.

Marini, F., Balestrieri, F., Bucci, R., Magry´, A. D., Magry´, A. L., & Marini, D. (2004). Supervised pattern recognition to authenticate Italian extra virgin olive oil varieties. *Chemometric and Intelligent Laboratory System*, 73(1), 85–93.

Massart, D.L.: , Kaufman.L. (1983). *The interpretation of analytical chemical data by the use of cluster analysis*. New York, Wiley.

Poulli, K.I., G.A. Mousdis, C.A. Georgiou, 2005. Classification of edible and lampante virgin olive oil based on synchronous fluorescence and total luminescence spectroscopy, *Anal. Chem. Acta.*, 542: 151-156.

Reid, L. M., O'Donnell, C. P., & Downey, G. (2006). Recent technological advances for the determination of food authenticity. *Trends in Food Science and Technology*, 17(7), 344–353. doi:10.1016/j.tifs.2006.01.006.

Savitzky, A.; Golay, M.J.E. (1964). "Smoothing and Differentiation of Data by Simplified Least Squares Procedures". *Analytical Chemistry* **36** (8): 1627–1639

Sayazo, A., Morales, M. T., & Aparicio, R. (2004). Detection of hazelnut oil in virgin olive oil by a spectrofluorimetric method. *European Food Research and Technology*, 218, 480–483.

Tay, A., Singh, R. K., Krishnan, S. S., & Gore, J. P. (2002). Authentication of olive oil adulterated with vegetable oils using Fourier transform infrared spectroscopy. *LWT-Food Science and Technology*, 35, 99–103.

Terouzi, W. ,Deluca,M., Gaetano, Oussama, A.(2011). A discriminant method for classification of Moroccan olive varieties by using direct FT-IR analysis of the mesocarp section .vibrational spectroscopy, 56, 123-128.

UNESCO (2007). "Biosphere Reserve Information" MAB Programme.

Résumé

L'huile d'olive représente l'un des produits stratégiques de la région euro-méditerranéenne. À l'instar des pays du bassin méditerranéen, le Maroc doit se plier aux exigences agronomiques et technologiques de son oléiculture.

L'arganier (*Arganiaspinosa*, Skeels L., Sapotacées) est un arbre endémique du sud-ouest marocain. En raison de la valeur élevée de ses huiles, la possibilité des pratiques illégales devrait être contrôlée.

La spectroscopie infrarouge à Transformée de Fourier (IRTF) est une technique puissante, et elle représente un outil d'analyse rapide et non destructive pour l'étude des huiles et graisses alimentaires, nécessitant peu ou pas de préparation des échantillons.

L'objectif de ce travail est d'explorer les potentialités de la spectroscopie IR et les méthodes chimiométriques pour prédire la qualité des huiles végétales

Dans ce travail, nous avons développé une nouvelle application basé sur la spectroscopie IRTF combinée avec des méthodes supervisées et des méthodes non-supervisées pour la classification des amandes d'argan par l'analyse hiérarchique ascendant. CHA a montré trois groupes distincts liés aux trois régions marocaines et a montré l'utilité de la spectroscopie infrarouge pour distinguer les origines des amandes d'argan. La PLS2 a été appliquée pour prédire l'origine géographique des amandes. L'autre partie de ce travail, elle a évalué l'efficacité de la spectroscopie infrarouge à transformée de Fourier (IRTF) à détecter l'adultération. Le premier axe porte sur la détection de l'adultération d'huile d'olive par l'huile de soja et l'huile de tournesol. La régression PLS et l'analyse en composante principale ACP ont été utilisée comme un outil analytique. Le deuxième axe porte sur la détection et la classification de l'adultération de l'huile d'argan par l'huile de soja et l'huile de tournesol.

Mots-clés : Spectroscopie Infrarouge, Chimiométrie, PLS, PCA, analyse discriminante, CPG, ATR, huile, FTIR, cluster analyses, Origine géographique, Amande d'argan