New insights into the Argan oil categories characterization: chemical descriptors, FTIR fingerprints, and chemometric approaches.

Kharbach, Mourad; Yu, H.; Kamal, Rabie; Barra, Issam; Marmouzi, Ilias; Cherrah, Yahia; Alaoui, Katim; Bouklouze, A.; Vander Heyden, Yvan

Published in:
Talanta

DOI:
10.1016/j.talanta.2020.122073

Publication date:
2021

New insights into the Argan oil categories characterization: chemical descriptors, FTIR fingerprints, and chemometric approaches

Mourad Kharbach[a,b*], Huiwen Yu[c], Rabie Kamal[d], Issam Barra[e], Ilias Marmouzi[b], Yahia Cherrah[b], Katim Alaoui[d], Abdelaziz Bouklouze[b] and Yvan Vander Heyden[a*]

[a] *Department of Analytical Chemistry, Applied Chemometrics and Molecular Modelling, CePhaR, Vrije Universiteit Brussel (VUB), Laarbeeklaan 103, B-1090 Brussels, Belgium*

[b] *Bio-Pharmaceutical and Toxicological Analysis Research Team, Laboratory of Pharmacology and Toxicology, Faculty of Medicine and Pharmacy, University Mohammed V in Rabat- Morocco*

[c] *Department of Food Science, University of Copenhagen, Rolighedsvej 30, DK-1958, Frederiksberg C, Denmark*

[d] *Pharmacodynamics Research Team, Laboratory of Pharmacology and Toxicology, Faculty of Medicine and Pharmacy, University Mohammed V- Rabat- Morocco*

[e] *Center of Excellence in Soil and Fertilizer Research in Africa CESFRA, Mohammed 6 Polytechnic University UM6P, Benguerir, Morocco*

*Corresponding authors:

E-mail address: yvanvdh@vub.be (Yvan Vander Heyden), mourad.kharbach@vub.be (Mourad Kharbach)

HIGHLIGHTS

- Chemical characterization and quality evaluation of Argan oil categories.
- Both chemical profiling and FTIR fingerprints investigated for categorical classification.
- Important chemical descriptors and FTIR sub-intervals highlighted for categorical description.
- Argan oils were classified as Extra virgin, Virgin or Lower quality.
- Chemometric tools allowed classification and discrimination between Argan oil categories.

Abstract

The characterization of Argan oils to classify them in three categories ('Extra Virgin', 'Virgin' and 'Lower quality') was evaluated. A total of 120 Moroccan Argan oils samples from the Taroudant Argan forest was investigated. The free acidity, peroxide value, spectrophotometric indices (K_{232} and K_{270}), fatty acids, sterols, and tocopherol contents were assessed. The samples were also scanned by FTIR spectroscopy. The Principal Component Analysis (PCA) and four classification methods, Partial Least Squares Discriminant Analysis (PLS-DA), Soft Independent Modeling of Class Analogy (SIMCA), K-nearest Neighbors (KNN), and Support Vector Machines (SVM), were applied on both the chemical and spectral data. Besides the conventional chemical profiling, FTIR spectra were evaluated for their feasibility as a rapid non-invasive approach for classifying and predicting the oil quality categories.

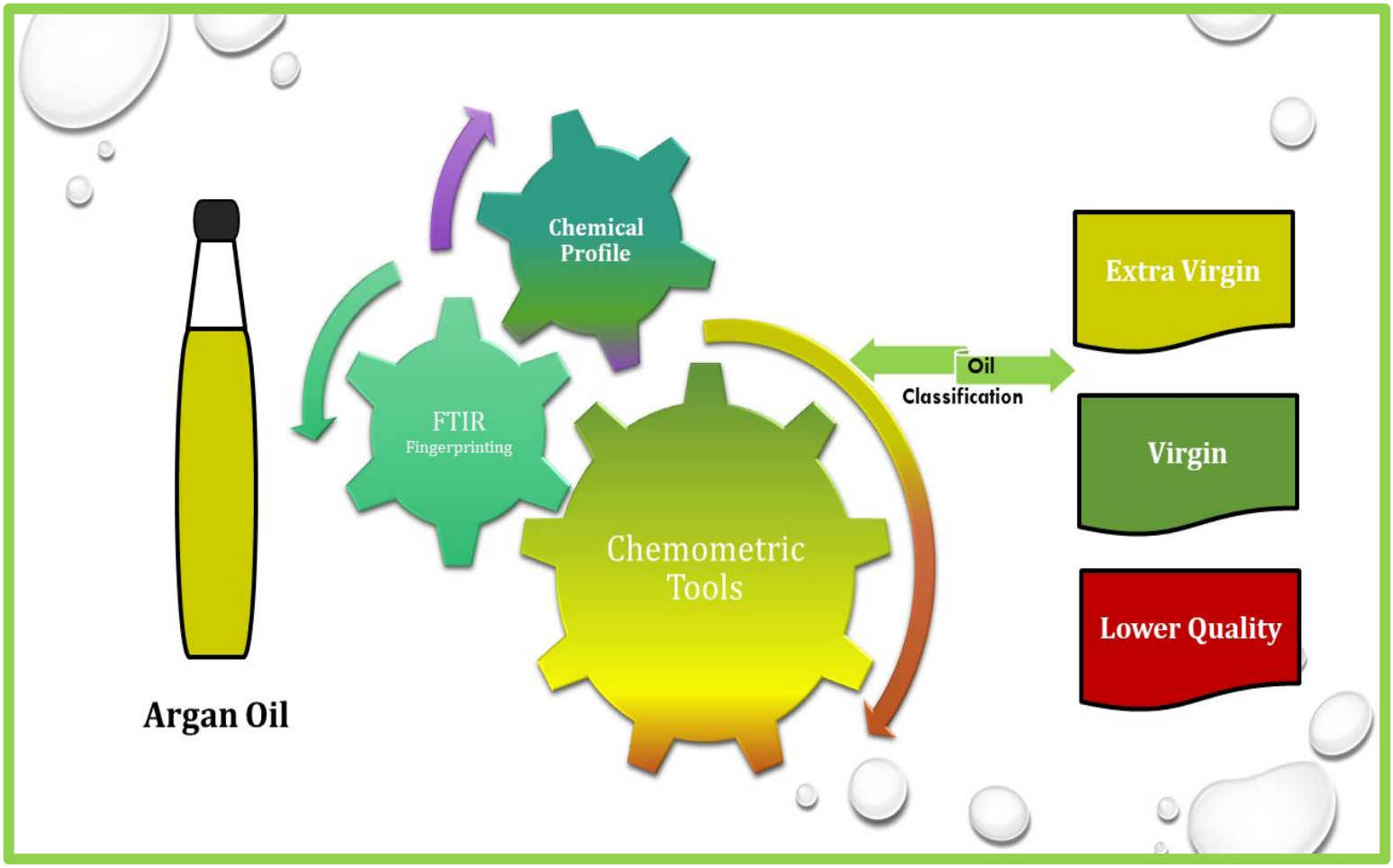
The most important variables for differentiating the oil categories were identified as K_{232}, peroxide value, ɣ-tocopherol, δ-tocopherol, acidity, stigma-8-22-dien-3β-ol, stearic acid ($C_{18:0}$) and linoleic acid ($C_{18:2}$) and could be used as quality indicators. Eight chemical descriptors or key features from the FTIR spectra (selected by interval-PLS) could also be established as indicators of quality and freshness of Argan oils.

Keywords: Argan oil; Oil traceability; Fourier Transform Infrared Spectroscopy; Quality evaluation; Classification tools; Chemometric analysis.

Chemical compounds studied in this paper

γ-Tocopherol (PubChem CID: 92729); δ-Tocopherol (PubChem CID: 92094); Linoleic acid (PubChem CID: 5280450); Oleic acid (PubChem CID: 445639); Palmitic acid (PubChem CID: 985); Stearic acid (PubChem CID: 5281); Schottenol (PubChem CID: 441837); Stigma-8–22-dien-3β-ol (PubChem CID: 5280794); Spinasterol (PubChem CID: 5281331); Δ−7-avenasterol (PubChem CID: 12795736).

Graphical abstract



1. Introduction

The Argan tree (*Argania spinosa L.*) is one of the oldest wild forest trees in Morocco and the world. It has been part of the Southwestern Moroccan regions and the Amazigh civilization since centuries. Argan production funds and ensures the subsistence of more than 3 million people (farmers, women's cooperatives, producers, sellers and speculators) [1]. The Argan oil production will be tripled in the upcoming years to reach 4000 tons. Argan oil is a natural product with a precious composition and properties destinated for dietary and/or cosmetic uses. The regular consumption of Argan oil has a positive effect on the prevention of certain diseases (heart and liver, cholesterol, and cancer) [2]. For many years, the scientific community has been investigating the Argan oil composition and its nutritional benefits [3].

Extra Virgin Argan oil (EVAO) is extracted from unroasted or roasted Argan fruits by mechanical or traditional extraction, under controlled conditions that do not cause compositional changes and preserve the oil properties. The Extra Virgin oil category has a specific composition, an acidity (expressed as oleic acid) not exceeding 0.8%, a peroxide value below 15 meqO2/kg, K_{270} inferior to 0.35 and other requirements, as shown in Table 1. The Virgin Argan oil (VAO) label is attributed to oil with specific acidity below 1.5%, peroxide value not exceeding 20 meqO_2/kg and K_{270} inferior to 0.35. Pure Argan oil quality is defined as oil with an acidity not exceeding 2.5%, peroxide value not exceeding 20 meqO_2/kg and K_{270} below 0.45. The lampante category (low quality Argan oil (LQAO)) concerns an oil with an acidity exceeding 2.5% [4, 5].

EVAO has an exclusive and highly different chemical composition, distinguishing it from other vegetable oils, which can be investigated for traceability and authentication [6]. The Argan oil composition variability is highly related to intrinsic and extrinsic factors, assembled into four groups: environmental (climate), agronomic (irrigation and fertilization), cultivation (harvesting period and maturity) and technological (storage and extraction process) factors [7-9]. These factors can extensively affect metabolites, as triglycerides, fatty acids, sterols, tocopherols, polyphenols and volatile compounds.

Nowadays, one of the major problems in the agricultural–food industry is to develop objective tools, which determine and ensure the quality and traceability of raw materials as well as of ~~finished~~ derived products. The determination and certification of product quality represents a form of protection for the consumer and a means of maintaining credibility and prestige for the food producer. An increasing interest for high-quality food products with a clear composition exists with consumers [10].

Classical chemical characterization and FTIR fingerprinting, associated with chemometric pattern recognition tools, demonstrated to provide satisfactory tools for recognizing the geographical origin of Argan oils [9]. Recently, fatty-acid profiling and UV-Visible spectra were also successfully applied in the EVAO characterization, geographic classification and extraction-process/kernel-type distinction [11]. Traditional analysis for food geographical classification, traceability or authentication has several drawbacks, i.e. being time- and solvent consuming, destructive, and requiring sample preparation. Spectroscopic techniques, which may overcome these hurdles, are fast, cheap and non-destructive. Spectroscopic methods, e.g. FTIR fingerprinting, combined with chemometric tools have a great popularity and progress rapidly in many fields [12]. FTIR spectra are successfully applied in quality control, authentication and food traceability, such as for olive oil [13], milk [14], meat [15], and Argan oil [16].

Chemometric tools are crucial mathematical approaches to extract useful information from chemical data of food samples in order to link with specific purposes, such as geographical classification, traceability and/or quality control. Sensory analysis by targeted and untargeted chromatography (GC), associated to ion mobility spectrometry (IMS), with chemometrics handling was applied for olive oil category classification (i.e. extra virgin olive oil, virgin, and lampante) [17, 18]. Phenol and polyphenol profiles combined with chemometric analysis allowed classifying red wines according to their geographic origin, grape variety and vintage [19].

The current work reports a search for reliable chemical indicators or FTIR sub-intervals (selected by interval-PLS) for Argan oil quality characterization and category distinction. Discrimination of high-quality extra virgin Argan oils from virgin and lower quality Argan oils by means of multivariate data analysis based on either chemical profiling or FTIR spectra, was also established.

2. Material and methods

2.1 Samples and storage conditions

In Southwestern Morocco, more than 46% of the Argan fruits production is from the Taroudant region. This region consists predominantly of Argan trees forest and produces high quantities of Argan fruits. The actual study focused on the Argan oil quality from that region resulting from three successive harvesting periods, 2015-2017 (avoiding the geographical

origin effect). A strategy of sample assemblage was implemented in order to extend the study over the different edible Argan oil categories. The sampled oils have same properties as those delivered to the consumer. In a collaboration with the women's cooperatives from Taroudant, healthy Argan fruits were collected to produce high-quality extra virgin Argan oils (25 samples) with Protected Geographical Indication (PGI). The roasted kernels were crushed using mechanical extraction, to produce edible EVAO (sample codes 1-25). The rest of the edible Argan oils (95 samples) was purchased directly from the different markets and sellers (women's cooperatives, local and traditional markets) from the Taroudant region. They were extracted mechanically using roasted kernels (sample codes 26-120) (Table 2). The samples were transported to the laboratory, transferred in dark glass recipients, and stored at 4°C. Later their chemical composition and FTIR spectra were determined.

2.2 Physico-chemical quality indices

The free acidity (expressed as % oleic acid), peroxide value (PV, expressed in milli-equivalents active oxygen per kilogram oil ($meqO_2/kg$)), and UV absorbance coefficients at 270 and 232 nm (K_{270} and K_{232}), were determined according to the methodology described in the European Commission Regulation EEC/2568/91 [20].

2.3 Fatty acid determination

The fatty-acid (FA) composition was achieved via the methyl-trans-esterification of the Argan oil according to the procedure from the European Official Methods of Analysis [20]. The analytical conditions are described in more detail in Kharbach et al [9]. In brief, the FAMEs (fatty acid methyl-esters) quantification was carried out by a gas chromatograph (6890/Agilent Technologies Wilmington, DE, USA) coupled to a flame ionization detector (FID). A HP-88 capillary column (100 m x 0.25 mm ID x 0.20 µm film) (Agilent Technologies Spain, Madrid) was used for chromatographic separation. Helium was the carrier gas with a flow rate of 1 mL/min. The injector-, detector- and oven temperatures were set at 230, 250 and 210°C, respectively. The FAMEs were identified the relative retention times using of FAME standards (purchased from Sigma-Aldrich, Lyon, France). The amounts were expressed as percent (%) of total fatty acids (g/100 g oil). Analyses were done in triplicate.

2.4 Tocopherol composition

The α, β, γ- and δ-tocopherols in Argan oil were quantified using the AOCS Method Ce8-89 [21]. The procedure was also described in Kharbach et al [6]. An amount of Argan oil (10 mg) was diluted in 100 mL *n*-hexane and injected in an HPLC instrument (1100 series, Agilent Technologies, Waldbronn, Germany) coupled with a fluorescence detector (G1321A). The excitation and emission wavelengths were set at λ_{exc} = 290 nm, λ_{emis} = 330 nm. The column was a ChromSpher C18 (25 cm × 4.6 mm × 5.0 µm) (Varian, Middelburg, The Netherlands). The mobile phase was a mixture of acetonitrile/methanol (50:50, v/v, isocratic mode) with a flow rate of 1 mL/min and 20µL sample was injected. The run time was 40 min.

Tocopherols identification and quantification were made using tocopherol standards (α-, β-, γ-, and δ-tocopherols; purchased from Sigma Aldrich, Lyon, France) dissolved in n-hexane (10 mg/ml and 6 dilutions) and then their calibration lines were applied for quantification.

2.5 Sterol Analysis

Argan oil sterol composition was determined according to the official method of the NP ISO 12228:2002 [22]. Sterol separation was done by means of a gas chromatograph (6890-Agilent model Technologies, Wilmington, DE, USA) equipped with a flame ionization detector (FID) using a DB-5 capillary column (30 m x 0.25 mm ID x 0.25 µm film) (Agilent) following the conditions previously described in Kharbach, et al. [9]. Briefly, the column temperature was maintained at 280 °C, while the injector and detector temperatures were set at 290 °C. The carrier gas was helium, with a flow rate of 2 mL/min. The individual sterols were expressed relative (%) to the total sterol amount, which was expressed as mg/100g Argan oil. Sterol composition was determined based on retention-time comparisons with standards (Sigma Aldrich, Lyon, France) and their calibration curves were used for quantification.

2.6 FTIR measurements

A Fourier transform infrared spectroscopy equipment (Bruker Optics, Ettlingen, Germany), coupled with an attenuated total reflectance (ATR) accessory and a deuterated triglycine sulfate detector, was used. The spectra were measured in the range 4000-400 cm^{-1}. A scan average of twenty spectra was recorded for each sample at ambient conditions and used in further data analysis.

2.7 Chemometric analysis tools

Chemometric analysis tools may be used for quality evaluation, authentication and distinction between the Argan-oil categories. The chemical-composition results or the FTIR spectra are gathered in data matrices **X**, while a **y**-vector indicates the oil categories (categorical response). The data matrices **X** (chemical composition or FTIR spectra) do not contain equivalent information, which thus may require different data handling in order to get the proper outcome.

Principal component analysis (PCA) is an unsupervised technique that provides latent variables (named principal components, PCs) by linear transformation of the initial variables. The PCA exploratory approach decomposes the original data set to both scores and loading matrices. The scores on given PCs represent the new coordinates and provide information on the distribution (topological behavior) of the samples, while the loading plot shows represent the relationship between the original variables [23]. In the present study, PCA was applied to evaluate the clustering of samples according to the oil categories based either on their chemical profiling or FTIR fingerprints. Further, PCA locates possible outliers, prior to building classification models.

Pattern recognition tools, including both linear and non-linear classification techniques, were tested for the classification and quality control of the Argan oils.

Partial least squares–discriminant analysis (PLS-DA) is a linear supervised classification technique, which combines the potential of partial least squares (PLS) regression and linear discriminant analysis (LDA). PLS-DA modulates the classes, one versus all others, and separates these by linear discriminant models. This technique aims to find the latent variables which discriminate and maximize the covariance between classes [24, 25], in this case study, the oil categories .

Soft independent modelling of class analogy (SIMCA) is also linear supervised pattern recognition technique for class modelling. SIMCA uses the PCA decomposition power for each class separately [26]. The SIMCA classification model has the ability to assign new unknown Argan samples to either none one, or more oil categories.

Support vector machines (SVM) is non-linear classification technique, which finds a hyperplane that separates the predefined classes linearly. It is a kernel-based technique for classification, which uses support vectors to define decision boundaries discriminating the classes [27].

K-nearest neighbors (KNN) is a non-parametric ~~and linear~~ classification technique. The KNN classification model is based on calculated distances (e.g. Euclidean distance) between samples in order to assign them to a class [26]. The KNN classification tools were used to assign unknown Argan samples to the predefined oil categories.

Interval-PLS (iPLS) is an algorithm developed for sub-interval and key-feature selection. The FTIR spectrum was divided into 50 equal subintervals containing 50 variables. The PLS regression is carried out to each sub-interval and the features are selected with regard to prediction relevance. The new data set is composed of the selected sub-intervals [28]. The iPLS algorithm was applied to the FTIR data to select the important sub-intervals (wavenumbers) for oil-category characterization.

Both for the chemical (initially 22 chemical descriptors, later reduced to 8) and FTIR (2540 variables, reduced to 1000) data, the 117 samples were randomly divided in calibration and validation subsets (see Table 2) in order to build suitable calibration and prediction models. The training set consisted of 87 samples and the test set of 30 ~~samples~~. The classification models (PLS-DA, SIMCA, SVM and KNN) were built on auto-scaled chemical variables, and on standard normal variate (SNV) and mean-centered pretreated FTIR data. The Venetian-blinds cross validation procedure was used to optimize the models [29].

The performance of the models was evaluated and optimized in terms of their sensitivity, specificity, classification rate, root mean squared error of calibration (RMSEC), root mean squared error of prediction (RMSEP), root mean squared error of cross validation (RMSECV), coefficient of determination of model fitting (R^2) and of cross-validation (R^2_{CV}).

In preliminary data elaboration, chemical parameters (see Table 1) were submitted to one-way analysis of variance (ANOVA), followed by a Tuckey's post hoc multi-comparison test at 5% level for statistical differentiation. They were applied to study the significant differences in the chemical parameters between the different oil categories. The most discriminant key descriptors for oil category characterization were selected based on their highly significant differences in averages and on PCA exploration. The iPLS algorithm on the other hand, allowed seeking relevant sub-intervals (key features) from the FTIR spectra.

The pattern recognition tools were applied using a MATLAB version 17.a software (The Math-Works, Natick, MA, USA) and the integrated PLS-Toolbox version 8.1 (Eigenvector Research, Wenatchee, WA, USA).

3. Results and discussion

This study investigates the potential of chemical profiling versus FTIR fingerprinting for Argan-oil quality characterization and category discrimination. Conferring to the chemical profiling, samples 118-120 were largely different from a standard Argan oil and were probably adulterated, Table 1. In brief, those samples were rich in linoleic acid ($C_{18:2}$, 50–60% of the total fatty acids) and in oleic acid ($C_{18:1}$, 18–21%). The tocopherols were also different and characterized by higher amounts of α-tocopherol (419-478 mg/kg oil) and ɣ-tocopherol (297-330 mg/kg oil), while β-and δ-tocopherols were rather similar (33-39 mg/kg oil). Sterols were higher in campesterol (6.7-7.5%) and stigma-8-22-dien-3β-ol (7.0-7.5%) contents, while schottenol and spinasterol (14-20%) were inferior in amount. From the Alimentary codex of vegetable oils [30], and considering the results it could be suspected these oils were mixed with a high percentage of sunflower oil. The samples were eliminated from further data analysis and the correct authentication was not included here.

3.1 Argan oil and Moroccan regulations

The higher demand for Argan-oil products from local and international consumers push the producers to deliver high-quality oil. The evaluation of the quality is necessary for consumption and worldwide exportation and is a priority for economic sustainability. The quality certification is based on a number of parameters, which should be within the legislative limits established by the Moroccan Commission of Normalization [4]. This guideline specifies the maximal parameter levels for Argan oil in order to classify it in four categories: extra virgin, virgin, pure and lampante [4, 5]. The oil category depends on some extrinsic and intrinsic factors, such as quality of raw Argan material, environmental conditions, harvesting period, ripeness, extraction process ~~applied~~ and storage conditions [7, 8]. The methods for Argan oil evaluation could be distributed into two groups: i) methods adopted by the Moroccan Normalization guidelines (classical methods); ii) methods proposed by academic researchers, including more advanced and sophisticated techniques. These latter methods are not yet included in the official guidelines. In this study, the chemical profiling (classical methods) is compared to an FTIR-based approach (sophisticated method) for the purpose of oil-quality characterization.

3.2 Physicochemical quality characteristics

The quality parameters and the oxidative status, assessing free acidity (FA), peroxide value (PV), K_{232} and K_{270} (indicators of primary and secondary oxidation products, respectively), were evaluated. The results (acidity, PV, K_{232} and K_{270}) revealed statistically significant differences (Table 1).

Concerning the acidity values, the samples from the first category ranged from 0.25 to 0.34% oleic acid, which is within the acceptable norm (⩽0.8% oleic acid) for "extra virgin" quality, and evidenced their higher quality and freshness [4]. The second category varied between 1.20 and 1.25% oleic acid, which is within the acceptable range (⩽1.50% oleic acid) for "virgin oil" [4]. The third oil category exhibited the highest free acid index (1.55 to 2.66% oleic acid). Higher FA amounts for Argan oils might indicate inappropriate extraction and processing practices [7], storage or post-harvest conditions [8], and/or Argan fruit maturity (ripeness) [31]. The peroxide value is a parameter to indicate Argan oil oxidation. The Moroccan PV limits of "extra virgin" and "virgin" Argan oils were defined as ⩽15 meqO_2/kg and ⩽20meq O_2/kg, respectively [4, 5]. The first category samples ranged from 2.0 to 2.5 meqO_2/kg and were within the "extra-virgin" limits, while the samples from the second category were ranging from 16.6 to 17.3 meqO_2/kg. The third Argan oil category had the highest PV, ranging from 15.5 to 24.5 meqO_2/kg, indicating the oxidation of fatty acids. The PV increases with the storage conditions [32].

The autoxidation of the unsaturated fatty acids were measured at 232 and 270 nm. K_{232} and K_{270} were indicated as ⩽2.52 and ⩽0.35, for both the "extra virgin" and "virgin" Argan oil categories, respectively [4]. The first and the second categories studied exhibited values within the "extra virgin" and "virgin" Argan oils limits. The third group exhibited values exceeding the acceptable limits.

Consequently, according to the values of FA, PV, K_{232} and K_{270} the samples may be classified as "extra virgin", "virgin" and "Argan oil with lower quality", respectively. The FA, PV, K_{232} and K_{270} values increased significantly from the extra virgin category to the Argan oils with lower quality.

3.3 Fatty-acid changes

The fatty-acid composition was earlier already investigated for the characterization, determination and distinction of the geographical origin of Argan oils [11].

The fatty-acid composition was inspected in order to get a quality characterization and the oxidative stability profile. Significant differences were found between categories ($p < 0.05$) based on the individual fatty-acid, saturated (SFA), mono-unsaturated (MUFA) and poly-unsaturated fatty acids (PUFA) contents (Table 1). Oleic ($C_{18:0}$) and linoleic acid ($C_{18:2}$) are the main fatty acids in EVAO and VAO. Higher MUFA and PUFA fractions are seen in EVAO. Oleic acid ($C_{18:1}$) and MUFA were both varying from about 43% in LQAO to 48% in EVAO, while linoleic acid ($C_{18:2}$) and PUFA occurred in fractions of 29% in LQAO and 35% in EVAO samples. Both MUFA and PUFA profiles of VAO are situated between LQAO and EVAO. Environmental conditions and the extraction process might affect the oleic and linoleic contents [9, 33].

Palmitic ($C_{16:0}$), stearic ($C_{18:0}$), arachidic ($C_{20:0}$) and myristic acid ($C_{14:0}$) increased significantly from EVAO to LQAO samples. LQAO showed the highest fraction of SFA (23%) compared to 18% in EVAO. The same tendency is seen for palmitic ($C_{16:0}$) (15% vs 12%), stearic ($C_{18:0}$) (7.27% vs 5.62%), arachidic ($C_{20:0}$) (0.48% vs 0.29%), and myristic acid ($C_{14:0}$) (0.21% vs 0.15%). The VAO revealed a wide-ranging SFA profile, except for the stearic-acid content.

The MUFA/PUFA, PUFA/SFA, palmitic/stearic, oleic/palmitic and oleic/linoleic ratios were examined for nutritional evaluation. These parameters were introduced in olive-oil analysis to evaluate both the nutritional quality and oxidative stability [34]. The highest of these ratios provides information on the ripeness of the fruit (kernel) and the process quality that preserves the properties during the oil extraction [11]. The ratios of MUFA/PUFA and oleic ($C_{18:1}$)/linoleic ($C_{18:2}$) acids were highest in LQAO, followed by VAO and EVAO samples. The PUFA/SFA and oleic ($C_{18:1}$)/palmitic ($C_{16:0}$) ratios on the other hand were highest in EVAO.

A PUFA/SFA ratio above 1.5 is linked with health and dietary benefits [35]. The Argan oil consumption is correlated with health benefits including heart disease prevention [36]. The PUFA/SFA ratios exceed 1.5 for EVAO and VAO samples and represent a positive health effect and nutritional benefits, contrary of LQAO samples (below 1.5) . The oleic ($C_{18:1}$)/palmitic ($C_{16:0}$) ratio showed a similar behavior as the PUFA/SFA ratio.

3.4 Tocopherol composition

Tocopherol occurs in natural Argan oil in four forms, α-, β-, γ-, and δ-tocopherol, as natural antioxidants [33].

Both the individual and the total tocopherol contents show a significant difference ($p < 0.05$, Table 1) between the three categories and decrease with decreasing quality. The total tocopherol concentration ranged between 798–909 mg/kg in EVAO, 751–809 mg/kg in VAO (decrease of 7%), and 419–531 mg/kg in LQAO (decrease of 45%).

γ -tocopherol is the dominant tocopherol with an average content of 704 mg/kg in EVAO, which was 3% lower in VAO (682 mg/kg) and 41% in LQAO (416 mg/kg). The δ-tocopherol content was approximately 32% lower in VAO (71 mg/kg) relative to EVAO (104 mg/kg) and was 63% lower in LQAO (39 mg/kg). The α-tocopherol content ranged between 15 (LQAO) and 41 mg/kg (EVAO), with 35 mg/kg in VAO. Further, β-tocopherol was found in trace amounts. The tocopherol composition may indicate the oxidation or deterioration state.

The tocopherol content is affected (decreased or increased) by extraction or roasting processes [33], origin [6], ripeness [31], oxidation stability [8].

3.5 Sterol composition

Sterols form a major fraction of the unsaponifiable composition in Argan oil, which may indicate its authenticity [33]. Schottenol and spinasterol are the abundant individual sterols in Argan oil, followed by Δ-7-avenasterol, stigma-8-22-dien-3β-ol, and traces of campesterol and stigmasterol. The total sterol content and the individual sterols differ significantly between the three categories and decrease with decreasing oil quality ($p < 0.05$, Table 1).

EVAO is rich in total sterols (217 mg/100g) followed by the VAO and LQAO categories with a 12% and 26% lower amounts, respectively. Schottenol was the major sterol and relatively seen varies limitedly between the different categories, 45.5% (LQAO), 47.4% (VAO) and 48.7% (EVAO) of the total sterols, while spinasterol ranged between 37.1% (LQAO), 40.9% (VAO) and 43.4% (EVAO). The Δ-7-avenasterol makes up 4.9% in LQAO, 5.5% in VAO and 6.1% in EVAO. The stigma-8-22-dien-3β-ol percentages are 4% (LQAO), 5.1% (VAO) and 5.6% (EVAO) of the total sterol fraction. The contents of cholesterol and campesterol were highest in EVAO. Kernel roasting (time and/or temperature), extractions (traditional or mechanical), and origin may influence the sterol composition in Argan oil [33].

3.6 Pattern recognition the chemical descriptors

3.6.1 PCA exploration

The PCA analysis was performed, on the auto-scaled chemical data (22 chemical descriptors) of the 120 samples. The score plot of the two PCs (Fig.1A, 92% of variance explained) shows a clear distinction between the three oil categories. The separation in three groups on the score plot is influenced by significant variables. The loading plot (Fig.1B) indicates the chemical markers responsible to differentiate the oil categories. The EVAO quality was influenced by δ- and β-tocopherols, oleic ($C_{18:1}$), linolenic ($C_{18:3}$), Δ^7-avenasterol, and cholesterol. The VAO was distinguished based on PV ~~peroxide value~~, palmitic acid ($C_{16:0}$) and α-tocopherol; while LQAO was distinguished by acidity, arachidic acid ($C_{20:0}$), K_{270} and K_{232}. The biplot displayed the relationships between oils (samples) distribution and chemical descriptors (variables) (Fig.1S Supplementary materials).

3.6.2 Indicator variables for oil categories characterization

The twenty-two chemical variables were investigated to select power discriminant variables for oil characterization. One-way anova, following the PCA, was performed. The variables are arranged by their discrimination power and significance (Table 3). All F-values for the chemical descriptors were largely above the critical $F_{(0.95,\ 3,\ 120)}$ value = 3.08.

Based on the anova results, and the Q-residual and T^2-Hotelling (Fig.2S Supplementary materials) plot from PCA exploration, eight variables are selected as key features, i.e. K_{232}, PV, ɣ-tocopherol, δ-tocopherol, acidity, stigma-8-22-dien-3β-ol, stearic acid ($C_{18:0}$) and linoleic acid ($C_{18:2}$). These variables could be used to indicate and evaluate the freshness and for oil category identification, while they may also be applied as index of oxidative changes.

Box-and-whisker plots are drawn for a better visualization of the eight discriminating variables, for comparison and insight tendencies (Fig. 2). They are useful overview graphics indicated the distribution variability of the variables in the three oil categories through medians, quartiles (Q_{25} and Q_{75}), outliers, which describe the data profile and central tendency. Best discriminating between the three oil categories were K_{232}, δ-tocopherol and stigma-8-22-dien-3β-ol.

3.6.3 Classification models

PLS-DA, SIMCA, SVM and KNN were applied to construct models for the discrimination between the three categories. Table 2 indicates the number of samples selected in the training and prediction subsets for each oil category. The Venetian blinds cross-validation procedure was applied to select the best classification models from the auto-scaled chemical data. Table 4A shows the results for the best classification models for the chemical profiles. The results from the PLS-DA, SIMCA, SVM and KNN models showed a perfect performance (100% classification rate) on the both subsets (training and test sets). The SIMCA model could be selected for further use. SIMCA is a powerful class-modelling technique to test whether a sample is authentic or not, to assign an unknown sample to a class (oil category) or not (outlier). The SIMCA model is established on the PCA decomposition of each class separately and the class space is defined based on two criteria, Hotelling T^2 and Q residuals. The SIMCA model based on chemical profiling is a suitable model for oil category classification and could be recommended.

3.7 Pattern recognition using FTIR fingerprinting

3.7.1 FTIR measurements

Fig. 3A shows the FTIR spectra between 4000 and 400 cm^{-1}. The oil spectrum is characterized by several bands (peaks), mainly attributed to the presence of oleic and linoleic fatty acids, triglycerides and peroxides. The FTIR spectrum acquired from the three Argan oil categories was similar but with different intensities (absorbances).

The spectra are characterized by the following absorbance peaks; bands at 2924 and 2852 cm^{-1} attributed to the symmetrical and asymmetrical stretching of aliphatic C–H in CH_2 and CH_3 groups; bands at 1743 cm^{-1} from ester carbonyl (C=O) group stretching vibrations (triglycerides); bands at 1463 and 1377 cm^{-1} attributed to CH_2 and CH_3 aliphatic groups (scissoring vibrations); a group of bands at 1238, 1163, 1114 and 1099 cm^{-1} obtained from the ester groups stretching vibrations (–C–O); and bands at 721 cm^{-1} related to the CH_2 rocking vibration of *cis*-di-substituted olefins [9, 37]. The FTIR spectra share two fingerprint regions, between 3029-2767 cm^{-1} and 1785-400 cm^{-1}. The first region is attributed to stretching vibrations (peroxides) and some variation in fatty acids content (secondary oxidation products). The second region is related to the conjugated bonds from triglycerides and bending vibrations of carbonyl aliphatic groups (CH_2 and CH_3) [38].

3.7.2 PCA exploration

PCA was performed after SNV followed by mean-centering pretreatment of the FTIR spectra. The first two PCs explained about 90% of the total variance. Fig.3B. shows the PCA score plot. A clear separation between the three Argan oil categories is seen. ~~Indeed, the samples were clustered into three groups; thus,~~ The FTIR results provide a similar grouping as the chemical data.

The data was reduced by eliminating redundant variables and selecting characteristic sub-intervals for oil category characterization.

3.7.3 Feature Selection based on interval-PLS

The FTIR dataset, 4000-400 cm^{-1} (2540 wavelengths) was subjected to interval-PLS to select key variables or region of interests (sub-intervals) thus eliminating redundant information. One thousand variables are extracted, including eight sub-intervals (3857-3786; 3457-3361; 3289-3149; 3007-2936; 2793-2653; 1802-1732; 1660-1235; 1164-882 cm^{-1}). The important variables contributing to the Argan oil category characterization are selected and further investigated.

3.7.4 Classification models

The classification models (PLS-DA, SIMCA, SVM and KNN) are build using the key features generated by iPLS. The classification results for the best models are summarized in Table 4B. The models of the four techniques provided good classification abilities: for the training set, resulting in 100% sensitivities and specificities. The prediction ability of the four constructed models was tested classifying 30 new samples (test set). Table 4B shows the good prediction abilities~~.~~ ~~–~~From the four models the SIMCA model based on sub-selected FTIR spectra is recommended for oil category classification or oil authentication.

Thus, to monitor the quality grade of the Argan oils, the FTIR fingerprints are fast, non-destructive, non-invasive and cheap compared the chemical profiling (classical approach).

4. Conclusion

The Argan-oil production sector is increasingly considered as a social and economic sustainable development item for the southwestern Moroccan regions. In that perspective, the Moroccan agricultural authority defined serval priorities, including Argan sector modernization, production increase, geographical origin protection and quality control enhancement.

The Argan-oil quality-control parameters and sensorial properties are defined in the Moroccan Normalization Guidelines. However, the Argan oil categories composition is not yet investigated. The latter purpose suggests focusing on the establishment of efficient analytical methods determining chemical profiles or FTIR fingerprints, and combining these with chemometric tools for a comprehensive characterization and classification of the Argan oil categories. Eight chemical descriptors, K_{232}, peroxide value, ɣ-tocopherol, δ-tocopherol, acidity, stigma-8-22-dien-3β-ol, stearic acid ($C_{18:0}$) and linoleic acid ($C_{18:2}$), were selected as important variables for Argan oil category characterization. On the other hand, FTIR fingerprints, with iPLS selected variable subsets, were proposed as fast alternative for the Argan oil category distinction.

Exploratory data analysis by PCA was carried out to evaluate whether similar samples occur as groups. Further, multivariate classification techniques, resulting in both linear (PLS-DA, SIMCA and KNN) and nonlinear models (SVM) were applied to the chemical descriptors or the FTIR sub-intervals. ~~All selected models~~ They were able to extract the required information and to classify the three Argan oil categories properly.

Conventional analytical procedures for chemical profiling to distinguish the Argan oil quality categories exhibit some drawbacks, such as high costs, time consuming, extensive sample preparation, and multiple cumbersome experimental assays. Alternatively, the FTIR technique is fast, non-destructive, non-invasive, accurate, requires no sample preparation and is a low-cost analytical technique. It may replace the conventional procedures to classify the Argan oils according to their quality.

Acknowledgements

The authors appreciate the support of the “Mohammed VI Foundation for Research and Protection of the Argan Tree”. They also acknowledge the financial support of Fonds Wetenschappelijk Onderzoek - Vlaanderen (FWO-Vlaanderen) (GO38816N), VLIR-UOS

(**Team project-VLIR 345 MA2017**), the Vrije Universiteit Brussel (VUB), and the Faculty of Medicine and Pharmacy-Rabat (FMPR).

Conflict of interest

The authors declare that they have no competing interests.

References

[1] D.F. Robinson, The Moroccan Argan Trade: Producer Networks and Human Bio-Geographies, 1st Edition ed., Taylor & Francis2020. https://doi.org/10.4324/9780367853556
[2] H.E. Monfalouti, D. Guillaume, C. Denhez, Z. Charrouf, Therapeutic potential of argan oil: a review, Journal of Pharmacy and Pharmacology 62(12) (2010) 1669-1675.
[3] Z. Charrouf, D. Guillaume, The argan oil project: going from utopia to reality in 20 years, OCL 25(2) (2018) D209. https://doi.org/10.1051/ocl/2018006
[4] Moroccain Normalisation Guidelines, Service de normalisation industrielle marocaine, Huile d 'argane. Spécifications. NM 8 (2003) 090.
[5] Z. Charrouf, D. Guillaume, Argan oil, the 35-years-of-research product, European Journal of Lipid Science and Technology 116(10) (2014) 1316-1321.
[6] M. Kharbach, R. Kamal, M.A. Mansouri, I. Marmouzi, J. Viaene, Y. Cherrah, K. Alaoui, J. Vercammen, A. Bouklouze, Y. Vander Heyden, Selected-ion flow-tube mass-spectrometry (SIFT-MS) fingerprinting versus chemical profiling for geographic traceability of Moroccan Argan oils, Food Chemistry 263 (2018) 8-17.
[7] B. Matthäus, D. Guillaume, S. Gharby, A. Haddad, H. Harhar, Z. Charrouf, Effect of processing on the quality of edible argan oil, Food Chemistry 120(2) (2010) 426-432.
[8] S. Gharby, H. Harhar, D. Guillaume, A. Haddad, B. Matthäus, Z. Charrouf, Oxidative stability of edible argan oil: A two-year study, LWT-Food Science and Technology 44(1) (2011) 1-8.
[9] M. Kharbach, R. Kamal, M. Bousrabat, M. Alaoui Mansouri, I. Barra, K. Alaoui, Y. Cherrah, Y. Vander Heyden, A. Bouklouze, Characterization and classification of PGI Moroccan Argan oils based on their FTIR fingerprints and chemical composition, Chemometrics and Intelligent Laboratory Systems 162 (2017) 182-190.
[10] F. Bianchi, M. Giannetto, M. Careri, Analytical systems and metrological traceability of measurement data in food control assessment, TrAC Trends in Analytical Chemistry 107 (2018) 142-150.
[11] M. Kharbach, R. Kamal, I. Marmouzi, I. Barra, Y. Cherrah, K. Alaoui, Y.Vander Heyden, A. Bouklouze, Fatty-acid profiling vs UV-Visible fingerprints for geographical classification of Moroccan Argan oils, Food Control 95 (2019) 95-105.
[12] L.C. Lee, C.-Y. Liong, A.A. Jemain, A contemporary review on Data Preprocessing (DP) practice strategy in ATR-FTIR spectrum, Chemometrics and Intelligent Laboratory Systems 163 (2017) 64-75.
[13] N. Sinelli, L. Cerretani, V. Di Egidio, A. Bendini, E. Casiraghi, Application of near (NIR) infrared and mid (MIR) infrared spectroscopy as a rapid tool to classify extra virgin olive oil on the basis of fruity attribute intensity, Food Research International 43(1) (2010) 369-375.
[14] E. Domingo, A.A. Tirelli, C.A. Nunes, M.C. Guerreiro, S.M. Pinto, Melamine detection in milk using vibrational spectroscopy and chemometrics analysis: A review, Food Research International 60 (2014) 131-139.
[15] A. Rohman, Y. Erwanto, Y.B.C. Man, Analysis of pork adulteration in beef meatball using Fourier transform infrared (FTIR) spectroscopy, Meat Science 88(1) (2011) 91-95.
[16] A. Oussama, F. Elabadi, O. Devos, Analysis of argan oil adulteration using infrared spectroscopy, Spectroscopy Letters 45(6) (2012) 458-463.

[17] B. Vega-Márquez, I. Nepomuceno-Chamorro, N. Jurado-Campos, C. Rubio-Escudero, Deep Learning Techniques to Improve the Performance of Olive Oil Classification, Frontiers in Chemistry 7 (2019).
[18] M. del Mar Contreras, N. Jurado-Campos, L. Arce, N. Arroyo-Manzanares, A robustness study of calibration models for olive oil classification: Targeted and non-targeted fingerprint approaches based on GC-IMS, Food Chemistry 288 (2019) 315-324.
[19] L. Jaitz, K. Siegl, R. Eder, G. Rak, L. Abranko, G. Koellensperger, S. Hann, LC–MS/MS analysis of phenols for classification of red wine according to geographic origin, grape variety and vintage, Food Chemistry 122(1) (2010) 366-372.
[20] European Commission Regulations, Commission regulation (EEC) No. 2568/91 of 11 July 1991 on the characteristics of olive oil and olive-residue oil and on the relevant methods of analysis., Official Journal of European Communities (1991) p. L248.
[21] AOCS, Official methods and recommended practises of the American Oil Chemist Society, , 5th Champaign, Illinois. (1998) 2-93.
[22] ISO-12228, (PN. EN. ISO), Animal and vegetable fats and oils determination of individual and total sterols contents—gas chromatographic method, (2002).
[23] R. Bro, A.K. Smilde, Principal component analysis, Analytical Methods 6(9) (2014) 2812-2831.
[24] D. Ballabio, V. Consonni, Classification tools in chemistry. Part 1: linear models. PLS-DA, Analytical Methods 5(16) (2013) 3790-3798.
[25] J. Viaene, Y. Vander Heyden, Introduction to Herbal Fingerprinting by Chromatography, Chemometrics in Chromatography, CRC Press, Taylor & Francis group, Boca Raton, 2018, pp 351-369.
[26] L.A. Berrueta, R.M. Alonso-Salces, K. Héberger, Supervised pattern recognition in food analysis, Journal of Chromatography A 1158(1) (2007) 196-214.
[27] G. Mountrakis, J. Im, C. Ogole, Support vector machines in remote sensing: A review, ISPRS Journal of Photogrammetry and Remote Sensing 66(3) (2011) 247-259.
[28] T. Mehmood, K.H. Liland, L. Snipen, S. Sæbø, A review of variable selection methods in partial least squares regression, Chemometrics and Intelligent Laboratory Systems 118 (2012) 62-69.
[29] F. Westad, F. Marini, Validation of chemometric models–a tutorial, 893 (2015) 14-24.
[30] Codex Alimentarius, Codex standard for named vegetable oils, Codex Stan 210 (1999) 1-13.
[31] H. Harhar, S. Gharby, B. Kartah, D. Pioch, D. Guillaume, Z. Charrouf, Effect of harvest date of Argania spinosa fruits on Argan oil quality, Industrial Crops and Products 56 (2014) 156-159.
[32] R. Aparicio, J. Harwood, Handbook of olive oil, Analysis and properties. 2nd ed Springer, New York (2013).
[33] M. Hilali, Z. Charrouf, A.E. Aziz Soulhi, L. Hachimi, D. Guillaume, Influence of origin and extraction method on argan oil physico-chemical characteristics and composition, Journal of Agricultural and Food Chemistry 53(6) (2005) 2081-2087.
[34] G. Beltrán, C. del Rio, S. Sánchez, L. Martínez, Influence of harvest date and crop yield on the fatty acid composition of virgin olive oils from cv. Picual, Journal of Agricultural and Food Chemistry 52(11) (2004) 3434-3440.
[35] F. Sakouhi, S. Harrabi, C. Absalon, K. Sbei, S. Boukhchina, H. Kallel, α-Tocopherol and fatty acids contents of some Tunisian table olives (*Olea europea* L.): Changes in their composition during ripening and processing, Food Chemistry 108(3) (2008) 833-839.
[36] A. El Abbassi, N. Khalid, H. Zbakh, A. Ahmad, Physicochemical characteristics, nutritional properties, and health benefits of argan oil: A review, Critical Reviews in Food Science and Nutrition 54(11) (2014) 1401-1414.
[37] G. Navarra, M. Cannas, M. D'Amico, D. Giacomazza, V. Militello, L. Vaccaro, M. Leone, Thermal oxidative process in extra-virgin olive oils studied by FTIR, rheology and time-resolved luminescence, Food Chemistry 126(3) (2011) 1226-1231.
[38] M. Lerma-García, E. Simó-Alfonso, A. Bendini, L. Cerretani, Rapid evaluation of oxidised fatty acid concentration in virgin olive oil using Fourier-transform infrared spectroscopy and multiple linear regression, Food Chemistry 124(2) (2011) 679-684.

Table 1. The Argan oils chemical composition (reported as mean values and standard deviations), and the EVAO regulation requirements.

Chemical parameters	EVAO regulation	Category									Deviating samples		
		EVAO (N=40)			VAO (N=40)			LQAO (N=37)			Sample 118	Sample 119	Sample 120
		Mean ± SD	Min	Max	Mean ± SD	Min	Max	Mean ± SD	Min	Max			
Quality indices													
Acidity (%)	≤ 0.8	0.28^{a} ± 0.02	0.25	0.34	1.22^{b} ± 0.01	1,20	1,25	**2.28^{c} ± 0.23**	1.55	2.66	2.05	2.14	2.25
Peroxide value ($meqO_2/kg$)	≤ 15	2.2^{a} ± 0.13	2.0	2.5	16.9^{b} ± 0.16	16,6	17,3	**20.0^{c} ± 1.10**	15.5	24.5	19.8	18.7	19.6
K_{232}	≤ 2.52	1.26^{a} ± 0.02	1.23	1.32	1.96^{b} ± 0.02	1,93	2,02	**2.80^{c} ± 0.02**	2.77	2.83	2.45	2.35	2.22
K_{270}	≤ 0.35	0.22^{a} ± 0.01	0.20	0.24	0.32^{b} ± 0.01	0,30	0,34	**0.49^{c} ± 0.05**	0.31	0.54	0.52	0.53	0.50
Fatty Acids (% of total fatty acid)													
$C_{14:0}$	≤ 0.20%	0.15^{a} ± 0.02	0.11	0.19	0.15^{a} ± 0.02	0.10	0.18	**0.21^{b} ± 0.02**	0.17	0.25	0.17	0.18	0.15
$C_{16:0}$	11.50 - 15.00%	11.85^{a} ± 0.30	11.22	12.31	13.79^{b} ± 0.35	12.99	14.56	**14.65^{c} ± 0.28**	13.96	15.11	7.26	6.82	6.75
$C_{18:0}$	4.30-7.20%	5.62^{b} ± 0.28	4.96	6.11	4.67^{a} ± 0.19	4.32	5.02	**7.27^{c} ± 0.26**	6.84	7.75	6.58	6.67	6.60
$C_{20:0}$	≤ 0.50%	0.29^{a} ± 0.03	0.15	0.34	0.35^{b} ± 0.04	0.28	0.44	**0.48^{c} ± 0.02**	0.44	0.51	0.51	0.49	0.46
$C_{18:1}$	43.10-49.00%	**47.68^{b} ± 0.62**	46.28	48.63	43.92^{a} ± 0.44	43.12	44.68	42.72^{a} ± 0.45	41.92	43.48	20.69	18.32	19.20
$C_{20:1}$	≤ 0.50%	**0.47^{c} ± 0.02**	0.44	0.52	0.43^{b} ± 0.05	0.24	0.52	0.29^{a} ± 0.05	0.21	0.36	0.30	0.35	0.33
$C_{18:2}$	29.30-36.00%	**35.37^{b} ± 0.63**	34.19	36.29	34.06^{b} ± 0.57	33.26	35.19	29.01^{a} ± 0.74	27.96	30.35	59.73	52.14	60.32
$C_{18:3}$	≤ 0.30%	**0.25^{c} ± 0.02**	0.21	0.29	0.18^{b} ± 0.04	0.10	0.26	0.11^{a} ± 0.03	0.05	0.16	0.15	0.13	0.08
SFA	--	17.91^{a} ± 0.46	16.63	18.62	18.95^{b} ± 0.40	17.87	19.57	**22.62^{c} ± 0.37**	21.77	23.34	14.52	14.16	13.96
MUFA	--	**47.76^{b} ± 0.62**	46.74	49.12	44.35^{a} ± 0.45	43.36	45.12	43.01^{a} ± 0.46	42.14	43.80	20.99	18.67	19.53
PUFA	--	**35.62^{b} ± 0.62**	34.55	36.53	34.24^{b} ± 0.59	33.40	35.42	29.12^{a} ± 0.74	28.05	30.41	59.88	52.27	60.40
MUFA/PUFA ratio	--	1.34^{b} ± 0.03	1.28	1.38	1.30^{a} ± 0.02	1.26	1.33	**1.48^{c} ± 0.04**	1.41	1.54	0.35	0.36	0.33
PUFA/SFA ratio	--	**1.99^{c} ± 0.05**	1.89	2.11	1.81^{b}± 0.04	1.72	1.89	1.29^{a} ± 0.04	1.22	1.37	4.12	3.69	4.33
$C_{16:0}/C_{18:0}$ ratio	--	2.11^{b} ± 0.11	1.87	2.31	**2.69^{c} ± 0.14**	2.71	3.21	2.02^{a} ± 0.09	1.83	2.17	1.10	1.02	1.02
$C_{18:1}/C_{16:0}$ ratio	--	**3.99^{c} ± 0.09**	3.81	4.18	3.19^{b} ± 0.07	3.07	3.19	2.92^{a} ± 0.06	2.80	3.05	2.85	2.69	2.84

$C_{18:1}/C_{18:2}$ ratio	--	$1.34^{b} \pm 0.03$	1.28	1.38	$1.29^{a} \pm 0.02$	1.25	1.33	**$1.47^{c} \pm 0.04$**	1.41	1.54	0.35	0.35	0.32
Tocopherols (mg/kg of oil)													
α-tocopherol	18-75	**$41.48^{c} \pm 3.91$**	34.26	53.16	$34.89^{b} \pm 2.71$	25.22	41.29	$14.94^{a} \pm 2.42$	9.22	18.26	435.65	419.33	477.52
β-tocopherol	1-5	**$2.98^{c} \pm 0.46$**	4.11	1.58	$1.60^{b} \pm 0.28$	1.12	2.15	$0.19^{a} \pm 0.13$	0	0.48	32.74	38.17	36.26
ɣ-tocopherol	640-810	**$703.90^{b} \pm 14.15$**	651.15	760.22	$681.56^{b} \pm 7.61$	657.30	696.41	$416.18^{a} \pm 23.27$	369.26	475.39	330.22	297.19	318.73
δ-tocopherol	54-110	**$103.84^{c} \pm 3.97$**	97.89	112.54	$70.97^{b} \pm 2.42$	0.28	0.44	$39.00^{a} \pm 3.43$	32.29	49.26	32.76	35.62	39.29
Total tocopherol	600-900	**$852.21^{c} \pm 14.98$**	798.06	909.03	$789.03^{b} \pm 9.74$	750.59	808.54	$470.30^{a} \pm 24.80$	419.22	531.06	831.37	790.31	934.8
Sterols (% of total sterols)													
Cholesterol	≤ 0.40	**$0.37^{c} \pm 0.02$**	0.34	0.40	$0.26^{b} \pm 0.02$	0.22	0.30	$0.17^{a} \pm 0.03$	0.05	0.22	0.53	0.56	0.61
Campesterol	≤ 0.40	**$0.38^{c} \pm 0.02$**	0.33	0.40	$0.18^{a} \pm 0.02$	0.15	0.22	$0.26^{b} \pm 0.03$	0.14	0.32	6.71	6.87	7.24
Δ^{7}-Avenasterol	4.00 - 7.00	**$6.12^{c} \pm 0.23$**	5.80	7.00	$5.45^{b} \pm 0.40$	3.31	5.77	$4.89^{a} \pm 0.45$	2.55	5.60	5.82	6.05	6.26
Stigma-8-22-dien-3β-ol	3.20 - 5.70	**$5.62^{c} \pm 0.10$**	5.45	5.86	$5.12^{b} \pm 0.14$	4.90	5.40	$4.01^{a} \pm 0.13$	3.79	4.25	6.95	7.53	7.14
Schottenol	44.00 - 49.00	**$48.70^{b} \pm 0.38$**	47.90	49.26	$47.43^{a,b} \pm 0.29$	47.08	48.05	$45.50^{a} \pm 0.31$	45.09	46.05	17.5	20.12	13.74
Spinasterol	34.00 - 44.00	**$43.41^{c} \pm 0.53$**	42.19	44.27	$40.87^{b} \pm 1.65$	31.15	42.11	$37.10^{a} \pm 0.44$	36.18	38.05	15.03	22.19	20.47
Total sterols (mg/100g of oil)	≤ 220 mg/100g	**$217.20^{c} \pm 12.41$**	209.32	219.10	$190.11^{b} \pm 11.13$	180.15	202.72	$161.34^{a} \pm 9.14$	152.13	170.50	360.40	416.20	385.50

The lowercase letters indicate significant differences in the same line (comparison between the studied categories, Tuckey's test).

EVAO, Extra Virgin Argan Oil; VAO, Virgin Argan Oil and LQAO, Lower Quality Argan Oil.

SFA, Saturated Fatty Acids; MUFA, Monounsaturated Fatty Acids and PUFA, Polyunsaturated Fatty Acids.

The bold numbers represent the highest mean of the three categories.

Table 2. Sample collection and their origin, coding and class categories; EVAO, extra virgin Argan oil; PGI, protected geographic indication; VAO, virgin Argan oil; LQAO, lower quality Argan oil; AAO, adulterated Argan oil.

Category (confirmed after analysis)	Number of samples	Collection time	Origin	Sample code	Training set	Prediction set
EVAO (PGI)	10	September 2015	Taroudant cooperatives	1-10	8	2
EVAO (PGI)	15	October 2016	Taroudant cooperatives	11-25	11	4
EVAO (PGI)	5	November 2016	Taroudant cooperatives	26-30	4	1
EVAO	3	March 2017	Taroudant local markets	31-33	2	1
EVAO	3	March 2017	Taroudant traditional markets	34-36	2	1
EVAO (PGI)	4	July 2017	Taroudant cooperatives	37-40	3	1
VAO	8	November 2015	Taroudant cooperatives	41-48	6	2
VAO	3	December 2015	Taroudant local markets	49-51	2	1
VAO	12	June 2016	Taroudant cooperatives	52-63	9	3
VAO	7	July 2016	Taroudant cooperatives	64-70	5	2
VAO	5	March 2017	Taroudant local markets	71-75	4	1
VAO	5	June 2017	Taroudant traditional markets	76-80	4	1
LQAO	6	December 2015	Taroudant cooperatives	81-86	4	2
LQAO	5	June 2016	Taroudant cooperatives	87-91	4	1
LQAO	5	July 2016	Taroudant local markets	92-96	4	1
LQAO	5	July 2016	Taroudant traditional markets	97-101	3	2
LQAO	6	October 2016	Taroudant local markets	102-107	4	2
LQAO	5	March 2017	Taroudant local markets	108-112	4	1
LQAO	5	June 2017	Taroudant traditional markets	113-117	4	1
AAO	3	June 2017	Taroudant traditional markets	118-120	-	-

Table 3. One-way anova results. The variables are arranged according to their discriminant power for oil category characterization.

Variables	*F-test*	*F-critical*	*p-value*
K_{232}	53840.70	3.09	< 0.01
Peroxide value	8320.36		< 0.01
ɣ-tocopherol	3749.34		< 0.01
δ-tocopherol	3636.47		< 0.01
Acidity	2205.14		< 0.01
Stigma-8-22-dien-3β-ol	1660.86		< 0.01
$C_{18:0}$	1092.92		< 0.01
$C_{18:2}$	1022.77		< 0.01
Schotenol	906.13		< 0.01
$C_{18:1}$	831.12		< 0.01
$C_{16:0}$	807.40		< 0.01
K_{270}	805.53		< 0.01
Campesterol	771.82		< 0.01
α-tocopherol	756.74		< 0.01
β-tocopherol	706.96		< 0.01
Cholesterol	698.86		< 0.01
Spinasterol	354.98		< 0.01
$C_{20:0}$	346.48		< 0.01
$C_{20:1}$	238.70		< 0.01
$C_{18:3}$	185.26		< 0.01
$C_{14:0}$	139.69		< 0.01
Δ^7-Avenasterol	107.83		< 0.01

Table 4. The best classification models cross-validation and prediction set results from (A) the chemical profiling (8 parameters), and (B) the FTIR data based on iPLS, of the 117 samples.

A) Eight chemical descriptors																				
													Sensitivity (%)				Specifity (%)			
	PLS-DA			SIMCA			SVM			KNN			PLS-DA	SIMCA	SVM	KNN	PLS-DA	SIMCA	SVM	KNN
Category	1	2	3	1	2	3	1	2	3	1	2	3								
Training set																				
1 EVAO	30	0	0	30	0	0	30	0	0	30	0	0	100	100	100	100	100	100	100	100
2 VAO	0	30	0	0	30	0	0	30	0	0	30	0	100	100	100	100	100	100	100	100
3 LQAO	0	0	27	0	0	27	0	0	27	0	0	27	100	100	100	100	100	100	100	100
Prediction set																				
1 EVAO	10	0	0	10	0	0	10	0	0	10	0	0	100	100	100	100	100	100	100	100
2 VAO	0	10	0	0	10	0	0	10	0	0	10	0	100	100	100	100	100	100	100	100
3 LQAO	0	0	10	0	0	10	0	0	10	0	0	10	100	100	100	100	100	100	100	100

B) FTIR data based on iPLS																				
													Sensitivity (%)				Specifity (%)			
	PLS-DA			SIMCA			SVM			KNN			PLS-DA	SIMCA	SVM	KNN	PLS-DA	SIMCA	SVM	KNN
Category	1	2	3	1	2	3	1	2	3	1	2	3								
Training set																				
1 EVAO	30	0	0	30	0	0	30	0	0	30	0	0	100	100	100	100	100	100	100	100
2 VAO	0	30	0	0	30	0	0	30	0	0	30	0	100	100	100	100	100	100	100	100
3 LQAO	0	0	27	0	0	27	0	0	27	0	0	27	100	100	100	100	100	100	100	100
Prediction set																				
1 EVAO	10	0	0	10	0	0	10	0	0	10	0	0	100	100	100	100	100	100	100	100
2 VAO	0	10	0	0	10	0	0	10	0	0	10	0	100	100	100	100	100	100	100	100
3 LQAO	0	0	10	0	0	10	0	0	10	0	0	10	100	100	100	100	100	100	100	100

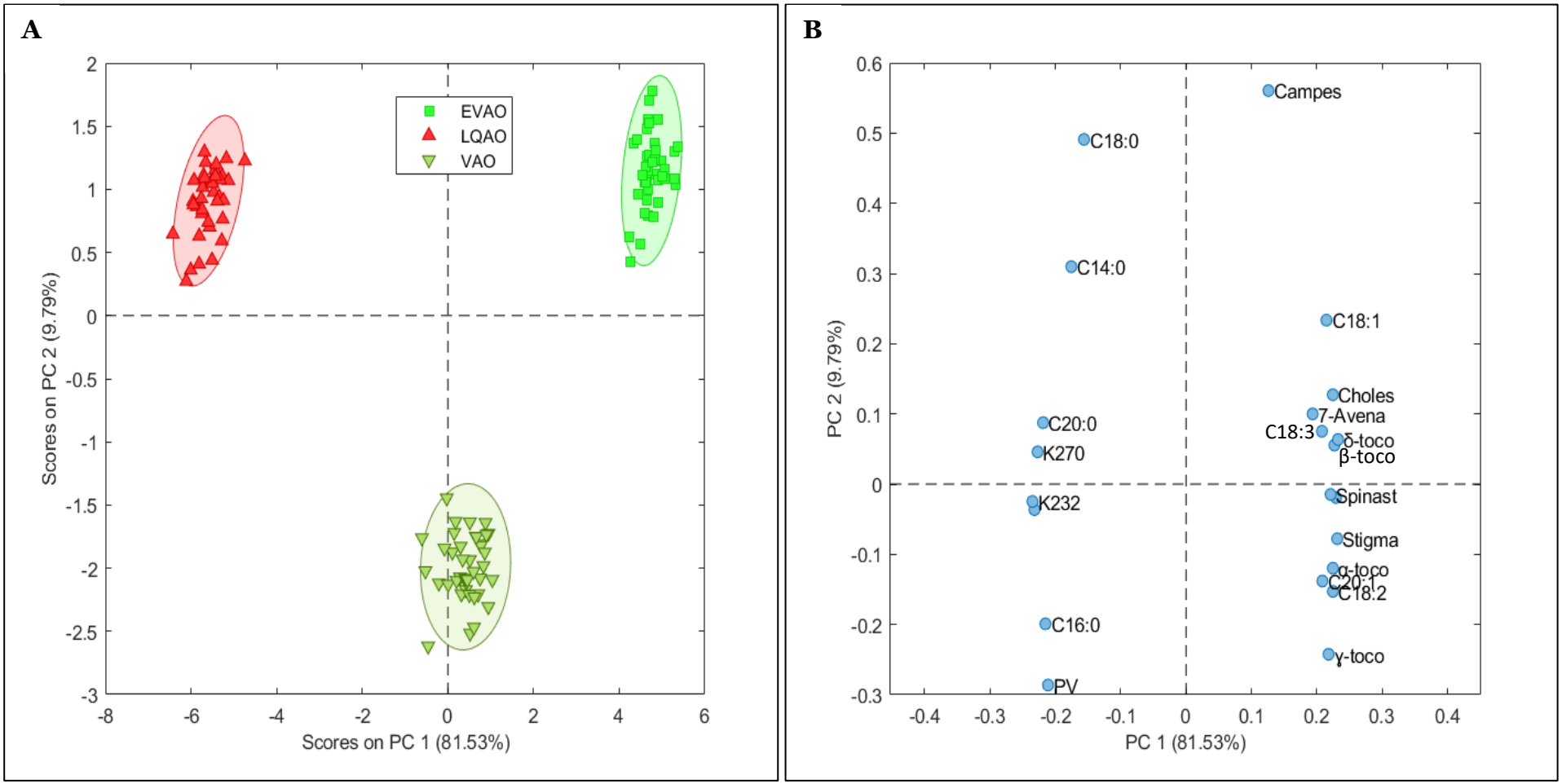


Fig. 1. (A) PC1-PC2 score plot, and (B) PC1-PC2 loading plot for the auto-scaled chemical profile.

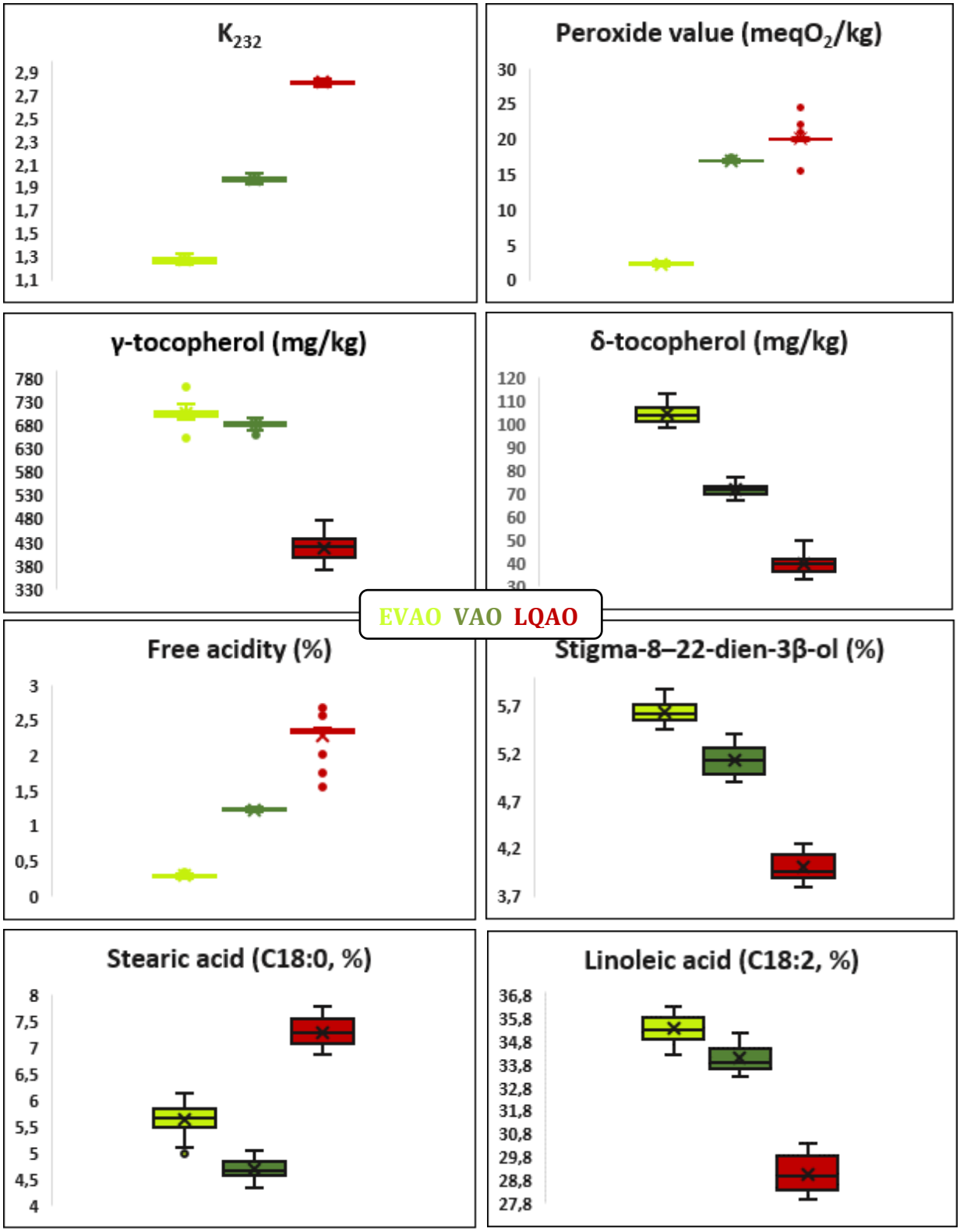


Fig. 2. Box-and-whisker plots, displaying the distribution of the best discriminating variables for the three Argan oil categories.

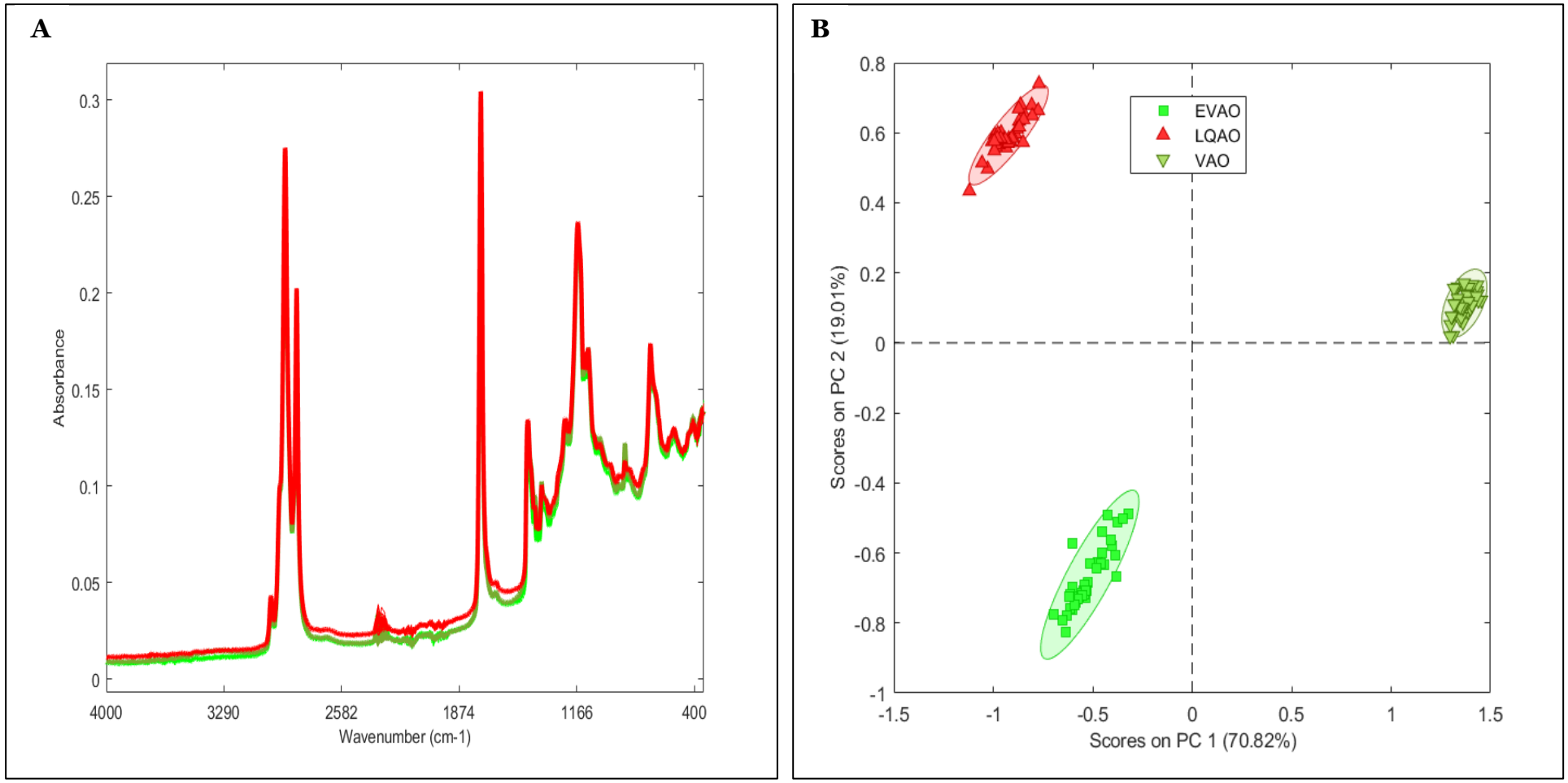


Fig. 3. (A) FTIR spectra of the Argan oil samples; (B) PC1-PC2 score plot for the SNV-treated and mean-centered FTIR spectra.

New insights into the Argan oil categories characterization: chemical descriptors, FTIR fingerprints, and chemometric approaches

Mourad Kharbach[a,b*], Huiwen Yu[c], Rabie Kamal[d], Issam Barra[e], Ilias Marmouzi[b], Yahia Cherrah[b], Katim Alaoui[d], Abdelaziz Bouklouze[b] and Yvan Vander Heyden[a*]

[a] *Department of Analytical Chemistry, Applied Chemometrics and Molecular Modelling, CePhaR, Vrije Universiteit Brussel (VUB), Laarbeeklaan 103, B-1090 Brussels, Belgium*

[b] *Bio-Pharmaceutical and Toxicological Analysis Research Team, Laboratory of Pharmacology and Toxicology, Faculty of Medicine and Pharmacy, University Mohammed V- Rabat- Morocco*

[c] *Department of Food Science, University of Copenhagen, Rolighedsvej 30, DK-1958, Frederiksberg C, Denmark Pharmacodynamics Research Team,*

[d] *Pharmacodynamy Research Team, Laboratory of Pharmacology and Toxicology, Faculty of Medicine and Pharmacy, University Mohammed V- Rabat- Morocco*

[e] *Center of Excellence in Soil and Fertilizer Research in Africa CESFRA, Mohamed 6 Polytechnic University UM6P, Benguerir, Morocco*

*Corresponding authors:

E-mail address: yvanvdh@vub.be (Yvan Vander Heyden), mourad.kharbach@vub.be (Mourad Kharbach)

Supplementary material

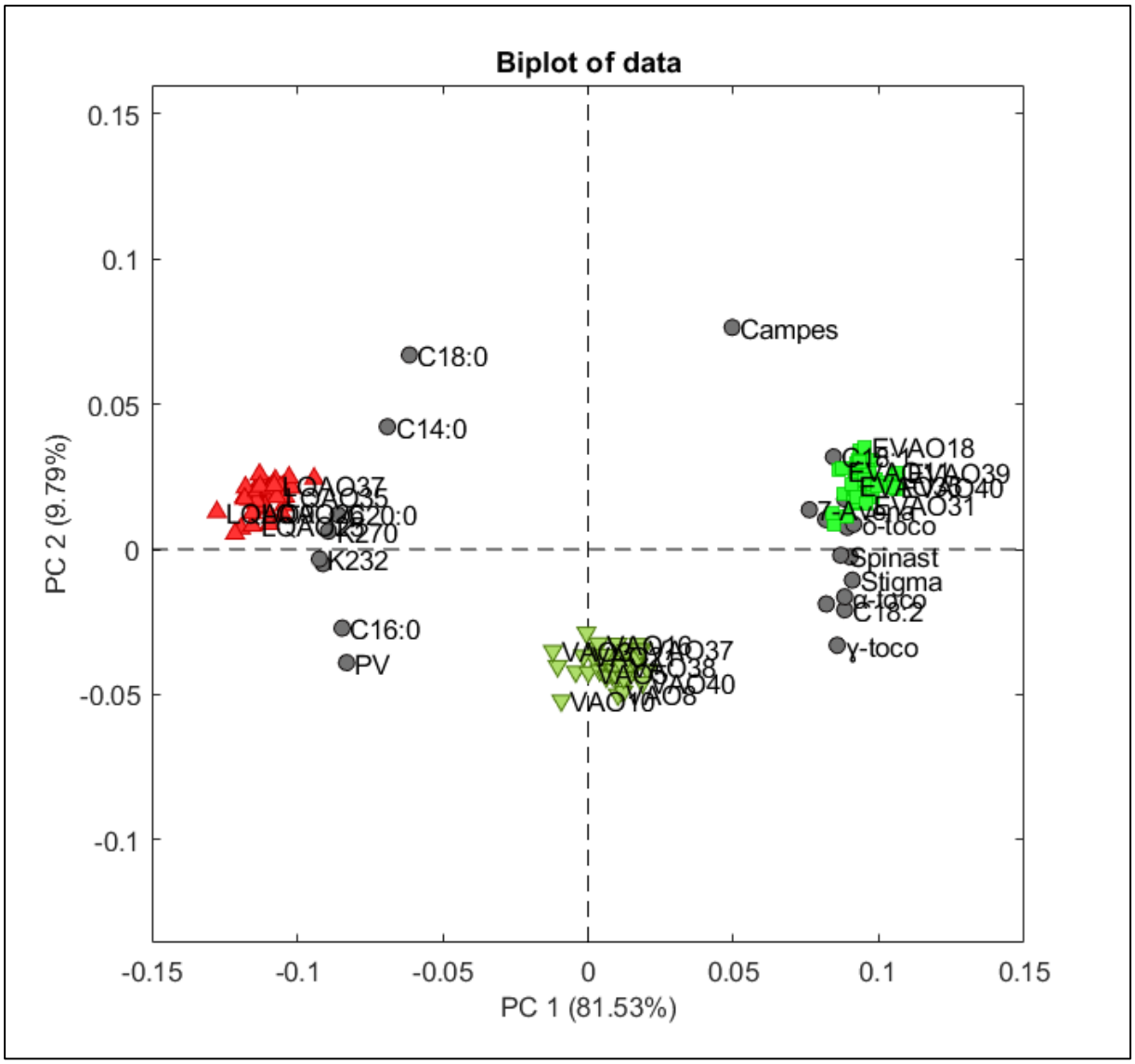


Fig. S1. PC1 vs PC2 PCA biplot of 117 samples and 22 chemical variables.

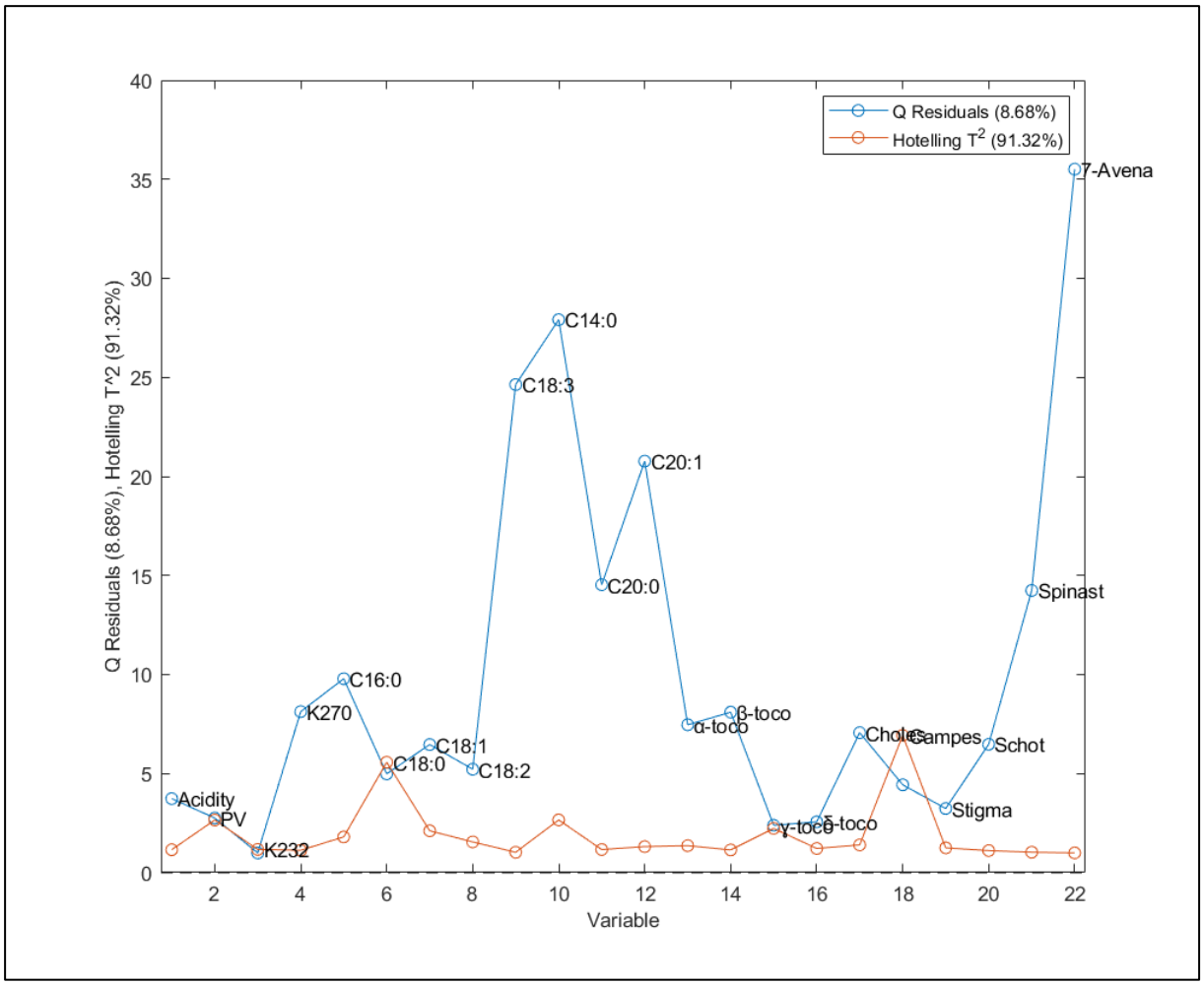


Fig. S2. Q residuals and T² Hotelling plot of the 22 chemical variables based on PCA data decomposition.