Universidad de Oviedo

POLYTECHNIC SCHOOL OF ENGINEERING OF GIJÓN

INDUSTRIAL CHEMICAL ENGINEERING DEGREE

ANALYTICAL CHEMISTRY AREA

OLIVE-OIL FRAUD DETECTION BY USING MINIATURIZED NEAR INFRARED DEVICES

D. TRAPIELLA FERNÁNDEZ, Jorge
TUTOR: Dra. SOLDADO CABEZUELO, Ana Belén
CO-TUTOR: Dr. FERRERO MARTÍN, Francisco Javier

DATE: June 2024

INDEX

FIGURE INDEX

TABLES INDEX

COMMON ABBREVIATURES

DLP- Digital Light Processing

DMD- Digital Micromirror microdevice

EU- European Union

EVOO- Extra Virgin Olive Oil

NIR(S)- Near Infrared (Spectroscopy)

OO- Olive Oil

PCA- Principal Component Analysis

PDO- Protected Designation of Origin

PGI- Protected Geographical Indication

RMS- Root Mean Square Error

SG- Savitzky Golay

SNV- Standard Normal Variate

UV/Vis- Ultraviolet/ visible

VOO- Virgin Olive Oil

1.- OBJECTIVE

The objective of this Project is the evaluation of miniaturized Near Infrared spectroscopy (NIRS) devices for the detection of olive-oil. NIRS is a fast, non-invasive, and cheap analytical technique to detect this type of frauds. The experiment carried out simulates an online process with a miniaturized device. To evaluate this NIR device, 160 samples with different oil compositions were analyzed by NIRS

It does not consist only in the mere evaluation of the mini spectrophotometer, but also the development of chemometric models using the appropriate chemometric tool to allow the clear detection of the adulterated olive oil samples.

Principal Component Analysis (PCA) will be used to develop different classification models and select those that allow the best contrast between olive oils and adulterated oils.

This proposed methodology can be implemented at industry level to allow the detection of possible frauds at the production site and in different stages of the production chain.

It also allows a non-specialized operator to use the automatized system and detect frauds at first sight without the need of any knowledge on the chemical field.

2.- SCOPE

The scope of the Project refers only to the development of a green analytical technique to detect frauds in olive oil. To attempt this work we have measured 160 oil samples (adulterated and non-adulterated with other seed oils) by means of a NIR miniaturized spectrophotometer, after that, different chemometric models to classify samples as adulterated or non-adulterated have been developed and evaluated. Graphical results based on Principal Components Analysis, that show that adulterated oils can be visually distinguished from pure olive oils, will be presented in this work.

3.- INTRODUCTION

Food fraud, defined as an intentional deception for economic gain using food, is a rapidly growing field of study because of increased health risk and cost awareness and, most importantly, increasing regulations. The purpose of enhancing production and benefits by minimizing costs often results in low quality products that can even be inimical [1], [2].

Most common adulterated products are those who meet two conditions: being high demanded, such as milk, and having an elevated market cost, such as honey, saffron or olive oil. Studies show that almost 25% of food frauds that are detected involve adulterations related to oils [3].

Olive oil, being one of the most valuable products in the Mediterranean area, is not only an expensive product, but has a very changing market due to fluctuating productions and prices, mainly motivated by strong regulation and the effect of climate change [4]. During the last three years, olive oil price has experienced an increasing tendency in the European Union, as shown in Figure 3.1 (a).

Figure 3.1 (b) shows that this inflation rate becomes even more accentuated in the case of Spain, being well over the European Union average and becoming the third country with the highest inflation rate in January 2024.

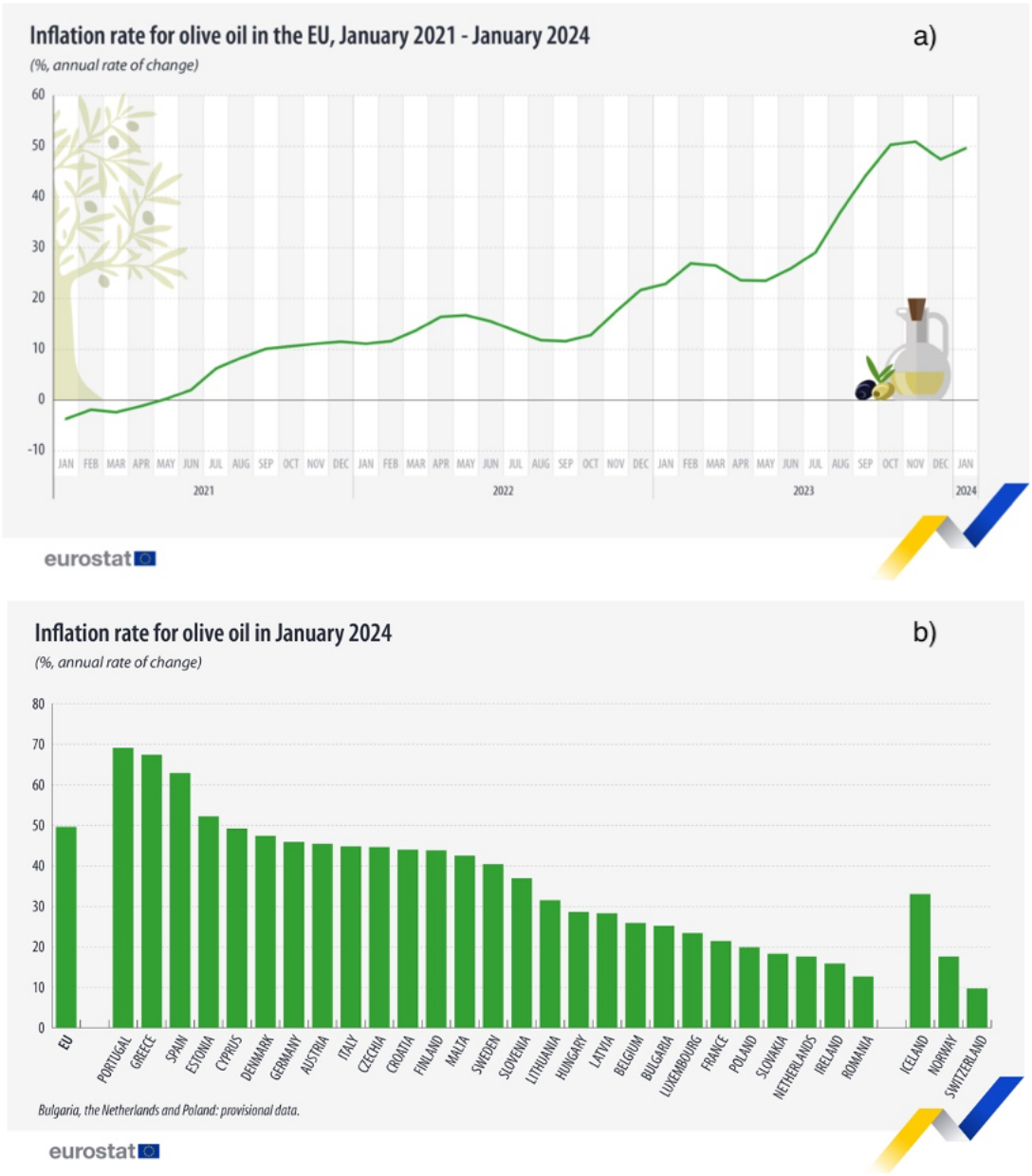

Figure 3.1- a) Inflation rate for olive oil in the EU, January 2021 – January 2024. b) Inflation rate for olive oil in January 2024 per country. Taken from Eurostat.

Lowering high production costs to increase benefits while maintaining commercialization prices often involves committing food fraud. Most common olive oil frauds include mislabeling, untrue origin, dilution, counterfeiting and substitution, being the latter half of the total frauds detected [5].

These frauds not only cause economic profits for the fraudster, but also may cause severe health and safety problems. The most famous was the episode related to the Spanish toxic oil syndrome (TOS) in 1981. Rapeseed oil fraudulently sold as olive oil caused an outbreak of an unknown disease. Approximately 300 people died because of TOS, a great number developed chronic disease and 20,000 people were affected in total, developing conditions such as incapacitating myalgias, marked peripheral eosinophilia and pulmonary infiltrates. A more recent incident was a health alert in Extremadura in March 2023 referred to the detection of more than 60,000 liters of adulterated oil being sold [6], [7].

Due to these incidents, the European Union has adopted strong measures to commercialize olive oil, which include the control to avoid adulterations by analytical methods. Some of these include acid base titration, UV absorption or gas chromatography (GC) [8]. These methods require lengthy sample preparation times, expensive equipment and may lack sensitivity to certain oil blends [9].

Over the last years, many alternative methods have ensured acceptable results to detect oil frauds, for example, fluorescence spectroscopy [10], ultra-high precision liquid chromatography (UHPLC) [11] or Raman spectroscopy [12]. However, most laboratory procedures require specialized personnel and high-cost techniques.

Nowadays, investigation leads to a non-invasive, fast, or, if possible, real-time detection and affordable technique. One of these techniques that fills these requirements is near-infrared spectroscopy (NIRS). Studies have already shown that NIRS is a valid analytical method to solve this problem [13], [14]. Although standard instrumentation may not be able to target directly olive oil adulterations, an instrumental proposal has been found that, with the use of chemometric models, can effectively detect this fraud [15].

3.1.- ANALYTICAL METHODOLOGIES TO DETECT ADULTERATED OILS

This section intends to explain the different methods for the detection of oil-related frauds mentioned in the previous section. Firstly, some of the methods proposed by the European Union will be explained and secondly, alternative methods will be analyzed.

The European Union, on its amending regulation *Reg. (EU) 2022/2104*, establishes the methods of analysis to be followed to determine characteristics of oils. Extra virgin olive oil (EVOO) must meet seven quality criteria and eight purity criteria for it to be commercialized under the declared category [8].

Among those, probably the simplest one is acid base titration to check free acidity in the oil, that is measured as oleic acid percentage. The principle of an acid base titration is adding a known volume of a basic substance (potassium hydroxide or sodium hydroxide) to the oil until the point of equivalence is reached, where the indicator changes the color of the solution [16]. The main drawback of this method is that it is destructive and that the point of equivalence may depend on the criteria of the operator [17].

Another method proposed by the EU is the use of gas chromatography (GC) to determine fatty acids. Before GC analysis a derivatization procedure to transform fatty acids to volatile fatty acids methyl esters (FAME) is required. This chromatographic methodology is based on the separation of different FAMEs by their different interaction with a carrier gas and a stationary phase. For the analysis of oils, the column used is often in the range of 8 to 12 meters length and 0.25 and 0.32 millimeters of internal diameter. The stationary phase is fused silica, and the detector used a flame ionization detector (FID). Figure 3.2 shows the chromatographic profile (GC-FID) of olive pomace oil, which identifies 24 FAMEs.

This method can quantify many FAMEs per analysis but has a very high cost and requires sample preparation by transesterification with methanolic solutions of potassium hydroxide [18].

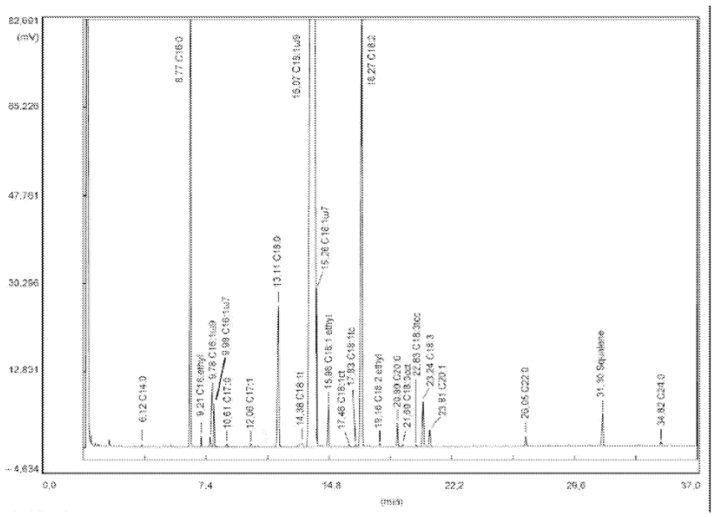

Figure 3.2- Chromatographic profile of olive pomace oil. Taken from [8].

Although these methods are destructive and sometimes require sample preparation while others do not, one characteristic they share is that they are time consuming and only target a specific analyte or group of analytes. This problem can cause to do many different analyses to ensure the quality of an oil [9].

This tedious series of methods has led to investigate on alternative methods to solve these inconveniences. For instance, Durán Merás [10] proposed fluorescence spectroscopy to discriminate between different olive oils according both to crop season and variety with a 100% of correct identification [19].

HPLC is an alternative method to GC. It is a cheaper method as the carrier is a liquid and requires less sample preparation [11]. Hayakawa [20] showed a simple method to target two compounds (squalene and tyrosol) and detect olive oil

frauds. It improves costs and time but does not solve some aspects such as sample destruction.

Routine analytical methods nowadays need three features: being fast, non-invasive and being implemented on an online process. This allows that analytical controls do not affect the chain of production and no product is destroyed. EU regulations, for example, require up to 10% of the oil produced to be given for analysis, which result in big losses for the producer [9]. Another desired feature is that multiple analytes can be targeted in a single procedure.

The two technologies that fit this description are Raman spectroscopy and NIR spectroscopy. The first one has been reviewed by de Lima [12] offering good results. However, this study will focus on the use of NIR, that has already been proven to be an effective method for the detection of oil adulterations [13], [14], [15].

NIRS has all the advantages stated before: it is fast, non-invasive and can be online. This study is centered in the detection of food frauds, although a method with these characteristics has also other benefits. For example, producers can commercialize the whole production, without the need of giving part for laboratory analyses. This technology could also be used to obtain information about the quality of the oil for internal controls. This will reduce costs as the quality and, consequently, the price of the oil can be established beforehand without the need of sending samples to a specialized laboratory.

To conclude this section, NIRS analyses offers much better results compared to other methods, as it does not only have benefits for control entities, but also for every subject involved in this industry.

4.- NIRS TECHNOLOGY

4.1.- THEORETICAL BASIS OF NIR ABSORPTION

Infrared radiation was discovered by William Herschel in 1800, while researching on modernizing the construction of telescopes. When using a prism and a thermometer to observe the thermal effects of radiation among the different ranges of the solar spectrum, he noticed that placing the thermometer outside the red border of the visible spectrum indicated an increase in temperature. This showed that the visible limits of the spectrum do not coincide with the limits of the solar spectrum. He described this unknown region of radiation as "dark heat", "invisible rays of sun" or "invisible thermometrical spectrum". Afterwards, it was renamed as infrared radiation [21].

The infrared region is the section of the electromagnetic spectrum that has a wavelength range from 780 to 25,000 nm. This region is commonly divided into three parts: medium (MIR) and far infrared radiation (FIR), that range from 2,500 to 25,000 nm and near infrared radiation (NIR), that is the closest section to the visible spectrum. NIR is considered to range from 780 to 2,500 nm [22].

Near infrared radiation spectroscopy (NIRS), has its basis on the observation of absorption bands caused by molecular vibration. Most of the molecules can be found at their fundamental vibrational energy levels at room temperature. Atoms that participate in chemical bonds are vibrating at a small range of frequency.

The two most common infrared active modes of vibration are those shown in Figure 4.1, known as stretching and bending. It is important to note that not all bonds absorb on the infrared, only organic compounds absorb NIRS radiation[23].

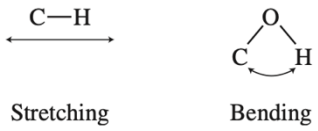

Figure 4.1- Common modes of vibration. Adapted from [23].

Radiation of a given frequency, if capable to give the molecule an energy equal to the difference between two vibrational levels, can be absorbed by the molecule and excite it to a higher vibrational level, as shown in Figure 4.2. This shows that only at a given wavelength will a bond absorb radiation. This relation between absorption and wavelength lead to the absorption spectra of a compound [24].

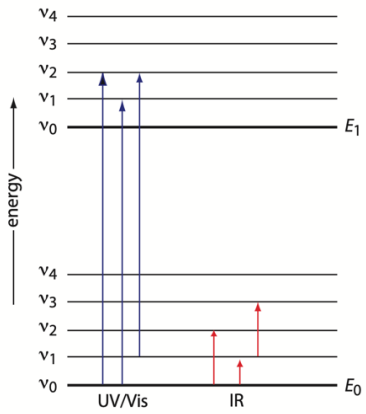

Figure 4.2- Absorption of different types of radiation. Taken from [19]

When NIRS radiation reaches a sample, the energy of this radiation is diminished as part of it is absorbed. It can be quantitatively described by two terms: transmittance and absorbance.

Transmittance refers to the ratio between the power that exits the sample and the one incident, as shown in equation 4.1.

$$T = \frac{P_T}{P_0} \tag{4.1}$$

Being T the transmittance, P_0 the incident power and P_T the power that exits the sample. Figure 4.3 shows this concept graphically. Note that P_0 is redefined as the power transmitted by a blank.

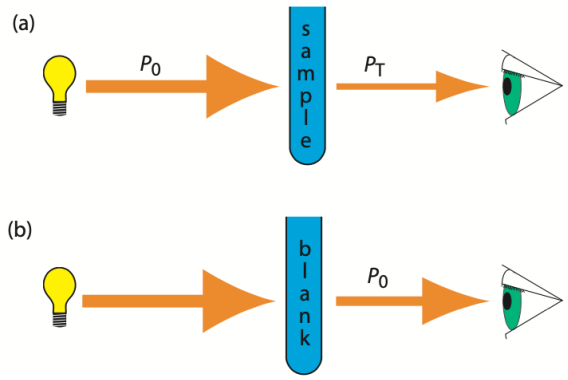

Figure 4.3- Transmittance. Taken from [19]

Absorbance is related to transmittance as shown in equation 4.2

$$A = -log\ T = -log \frac{P_T}{P_0} \tag{4.2}$$

Absorbance is a linear function of the concentration. This is shown by equation 4.3 known as Lambert-Beer law.

$$A = \varepsilon bc \tag{4.3}$$

A refers to the absorbance of the sample, ε is the molar absorptivity, that is characteristic to the compound, *b* is the pathlength and *c* is the concentration.

Then, absorbance values obtained on a sample can be related to the concentration of the existent species in the sample, if those absorb NIR [19], [23], [24].

4.2.- CHEMOMETRICS AND ITS APPLICATION TO NIRS TECHNOLOGY

The aim of this section is to explain the fundamentals of chemometrics, its application to NIRS technology and the tools to be used during the experiment.

Chemometrics can be defined as the science of relating chemical measurements made on a chemical system to a property of interest, such as concentration, through the application of mathematical or statistical methods [25].

This science emerged during the early 70s, when working with computers became more accessible for scientists. At first, it was only accessible to engineers and mathematicians. Jurs [26], Weiner [27] and Massart [28] published articles now considered as chemometric applications. However, it was not until 1972 that Wold published the first article containing the word chemometrics [29]. Wold joined Kowalski in 1974 to establish the International Chemometrics Society [30].

Extracting relevant information from NIR spectra is a challenge due to the complexity of spectral data and to different effects that can affect NIR spectra collection. These can be classified into three groups [31]. The first one is related to the instrument (signal-to-noise ratio, accuracy and precision of wavelength or signal linearity). The second one is related to the sample (homogeneity, density, texture, granulometry, stability to temperature changes...). The last group of errors is related to the human operations. Examples of this group are sample preparation, analyst errors or wrong procedures [31], [32].

Chemometric tools are employed to minimize these errors and to extract the relevant information included in spectra. Among all chemometric tools used for NIRS analysis, the most common group of methods is what is known as multivariate analysis [31].

Multivariate analysis can be defined as the set of statistical methods that consider more than one variable simultaneously for the analysis. Whereas classical univariate approaches underutilize the global data structure, losing a great amount of information, multivariate techniques allow exploitation of this data and the collecting of both qualitative and quantitative information through modelling. For NIRS technology, multivariate analysis is used to estimate any property of the sample through multiple spectral variables (absorbance values at different wavelengths) [33], [34], [35].

When light interacts with a different phase, many processes occur, such as refraction, reflection, and diffraction. These processes are often lumped together under the term "light scatter". To try to eliminate this effect, several chemometric techniques can be used. The most used are Multiplicative Scatter Correction (MSC) [36], orthogonal signal correction (OSC) [37] and Standard Normal Variate (SNV) [38].

SNV, introduced by Barnes [39], is the method to be used in this TFG during the chemometric treatment of the data. It allows the correction of the baseline caused by scattering. It centers and scales all spectra so that each of them has a mean of 0 and a standard deviation of 1. Equation 4.4 shows the mathematical expression:

$$x_{i,j}^{SNV} = \frac{(x_{i,j} - \bar{x}_i)}{\sqrt{\frac{\sum_{j=1}^{p}(x_{i,j} - \bar{x}_i)^2}{p-1}}} \qquad (4.4)$$

Where $x_{i,j}^{SNV}$ is the element of the transformed spectrum, $x_{i,j}$ is the corresponding original element of the spectrum *i* at variable *j*, $\bar{x}_i$ is the mean of the spectrum *i* and *p* is the number of variables or wavelengths in the spectrum [40].

MSC is a much more complex technique. However, research has shown that it gives the same results as SNV [41].

Another math pre-treatment often applied to NIR spectra are derivatives. First and second derivatives allow correcting the effect of peaks superposition and eliminates the displacement of the spectra baseline. The first order derivative minimizes or eliminates the baseline shifting, while the second order derivative minimizes or eliminates the shifting of the spectra that vary with the wavelength and, consequently, peak superposition. Derivatives with a higher order have no chemometric interest in NIR analysis [42].

Different methods are used for the calculation of derivatives. The most is Savitzky-Golay method [43]. Savitzky-Golay method was invented by Abraham Savitzky and Marcel Golay in 1974 [44]. It consists in taking a "window" of *n* points around each of the spectral data, construct a polynomial with them and replace the data with the point fitting that polynomial [45]. It is important when applying Savitzky- Golay method to select appropriately the number of points *n* to be in the "window". The more points are selected, the more noise is eliminated. However, an excessive number of points may eliminate peaks with relevant information [44].

Once spectral data obtained by NIRS have undergone correct math pre-treatments, chemometric models can be applied to obtain qualitative and quantitative information [31]. This research work intends to obtain qualitative information, applying chemometric techniques that allow distinction between olive oil and adulterated blends or seed oils, based on Principal Component Analysis (PCA). PCA is widely used in multidimensional analysis of data. It

identifies and reduces features that represent the original data in a lower multidimensional space, minimizing losses of information (data variability) [25].

4.3.- NIRS EQUIPMENT

Having explained before the flexibility of NIRS technology, this section intends to explain the different options referred to the equipment. In Figure 4.4 have been shown the basic components of a NIR spectrometer. The sample receives radiation from a light source. It will be absorbed at a certain wavelength, depending on the organic bonds of the samples. The ratio between the incident light and the one that exits the sample provides the quantitative information.

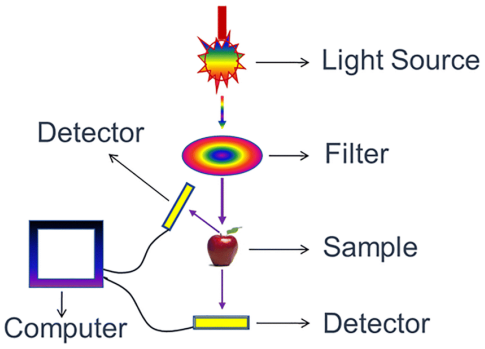

Figure 4.4- Schematic of a near infrared spectrophotometer. Taken from [46].

In NIR analysis, the light sources are required to emit in the 700-2500 nm region. The most common light source types are thermal radiation, light emitting diode (LED) and laser diode [22], [24]. Thermal radiation consists of passing an electric current through a tungsten filament to raise its temperature. For this purpose, tungsten halogen lamps are used [47]. LEDs and laser diodes are semiconductors that receive a voltage so that the excess energy is emitted as light. They have a more limited wavelength range than tungsten lamps. The most common LED is gallium arsenide (GaAs) and is used in communication and medical fields.

Other types of light source are solid state lasers, such as the neodymium-doped yttrium aluminum garnet (Nd:YAG) laser and titanium sapphire lasers, and supercontinuum light, which related to ultrashort light pulses introduced in nonlinear optical materials [24].

After the light is emitted, filters are used to disperse the incident light into different wavelengths. The most classical types of filters are prisms and diffraction gratings. Prisms are optical materials that simply discompose the incident light. Although its simplicity and low cost, they are normally inadequate for most NIRS analysis.

Nowadays, most NIR devices use diffraction gratings. Figure 4.5 shows a scheme of a diffraction grating.

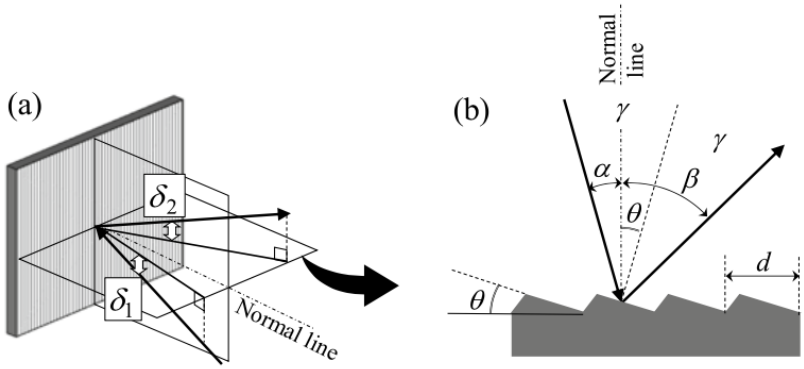

Figure 4.5- Diffraction grating. From [24]

When the light incises in the grating, a specific wavelength λ is diffracted at a specific angle θ. Normally, the separation of the angles can be small, so mirrors can be used to open these angles.

More advanced techniques involve the use of the Fourier Transform (FT). It is based on the measurement of the interference of two light beams (sample beam and reference beam) with a device called double beam interferometer [48].

Although it may not be relevant to the analytical problem to be solved, it is worth mentioning that there are filters named bandpass filters, that transmit light only of certain wavelengths. These include interference filters, variable filters, liquid crystal turnable filters (LCTF) and acousto-optic turnable filters (AOTFs).

Once the light is diffracted, the spectra is collected in the detector. NIRS detectors are made of semiconductors. Some examples are InGaAs (indium gallium arsenide), InAs (indium arsenide), InSb (indium antimonium); PbS (lead sulfide) and Si (silicon) [49]. Table 4.1 shows the different detection spectral range of different detectors.

Table 4.1- Classical NIRS detectors. From [50].

Detector	Spectral range of detection (nm)
InAs (indium arsenide)	1700-5700
InGaAs (indium gallium arsenide)	900-1700
InSb (indium antimonide)	1800-6800
MCT (mercury cadmium telluride)	1000-17000
PbS (lead sulfide)	1100-3000
PbSe (lead selenide)	1700-5500
PbTe (lead telluride)	1500-4500
Silicon	300-1100

Once the different parts of NIRS equipment have been explained, it is important to define the four analysis modes. This are transmittance, reflectance, transflectance and interactance [51]. These are shown in Figure 4.6.

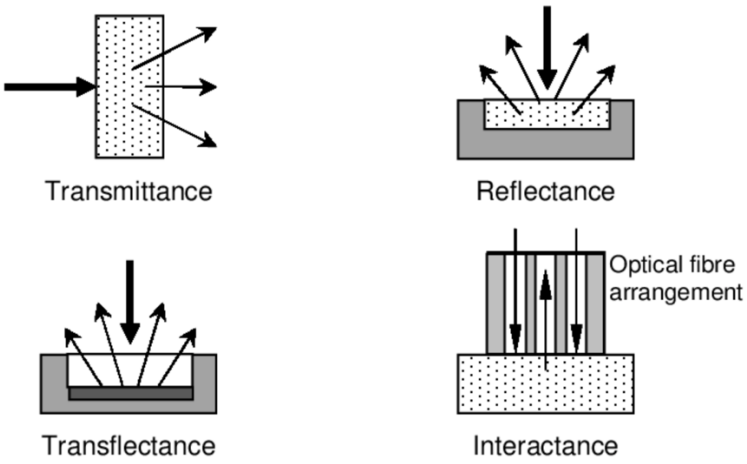

Figure 4.6- NIRS analysis modes. Taken from [52].

The two most common modes are transmittance and reflectance. In transmittance mode the detector is placed behind the sample, while in reflectance mode the detector is placed on the same plane as the sample [50]. The configuration of a spectrometer in transmittance and reflectance mode is shown in Figure 4.7.

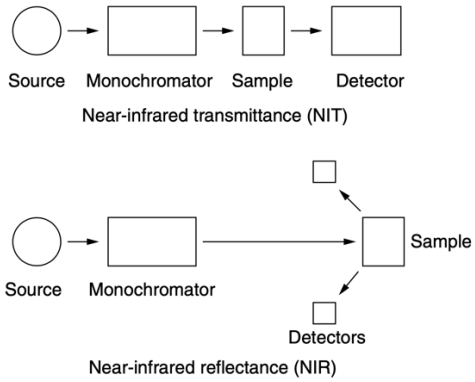

Figure 4.7- Transmittance and reflectance mode. Taken from [24].

Generally, laboratory equipment has better resolution being more expensive for the good quality of the pieces. Notwithstanding, investigation leads now to reduce the size of the spectrometer while maintaining the performance and lowering the price. Mini spectrometers are now manufactured with extremely

high quality that make NIRS technology adapted to be used on-site and implemented on industrial processes.

These portable sensors have been implemented in the food industry during the last years. Table 4.2 shows some examples of NIR handheld spectrometers that have been used in the food industry for the detection of frauds.

Table 4.2- Technical specifications of portable/handheld NIR devices. Adapted from [53].

Commercial name	Manufacturer	Light source	Wavelength selector	Detector
MicroPhazir	Thermo Fisher Scientific Inc.	Tungsten	MEMS	InGaAs
NIRONE	Spectral engines	Two tungsten vacuum lamps	MEMS	InGaAs
MicroNIR 1700	VIAVI Solutions	Two tungsten vacuum lamps	LVF	InGaAs
MicroNIR 2200	VIAVI Solutions	Two tungsten lamps	LVF	InGaAs
DLP NIRscan Nano	Texas Instruments Incorporated	Tungsten	LVF	InGaAs

Table 4.2 shows that the technologies used when miniaturizing the device show many similarities between them. The light sources come always from thermal radiation and are tungsten lamps. The detectors employed for handheld devices are always InGaAs, but there are different types of wavelength selector.

Micro-electro-mechanical systems (MEMS) are microdevices that include sensors, control circuits and mechanical structures, such as diffraction gratings using semiconductor materials [54]. Linear Variable Filters (LVF) are combinations of

the described filters before, combining bandpass filters (that allow each a certain wavelength range) to cover a wider spectral range [55].

This wide variety of analysis modes, light sources, filters, and detectors make NIRS technology to be extremely flexible and adaptable. Adding that to the advantages of this technology, such as being fast and non-destructive, NIRS is one of the most useful analytical methods for the resolution of many problems in different fields.

4.4.- USE OF NIRS FOR THE DETECTION OF FOOD FRAUDS

Many food products are legislated under a Protected Designation of Origin (PDO) or a Protected Geographical Indication (PGI), which establish systems of quality levels. The first one refers to products produced, processed, and prepared in a specific location and following a certain process, while the second only requires one of the stages of production, processing and preparation to take place in a given geographical area. These products are controlled by Regulating Councils, which guarantees the special characteristics the product has [22].

Over the last years, NIRS technology has emerged in this field for the detection of food frauds, as these products have usually high prices and limited production. This section intends to review the use of NIRS in different well-known products to detect possible adulterations.

A clear example is Iberian ham, which is one the most expensive luxury foods in Europe. The quality of the ham is influenced by the breed of the pig and the number of *montanera* days. *Montanera* is a stage in which the Iberian pigs feed on acorns and grass, available in the *dehesa* ecosystem [22], [56]. This feeding conditions highly influence the composition of the final product. Hernandez-Jiménez [56] showed that NIRS technology could classify correctly 100% of the

samples to its corresponding duration of the *montanera* using fatty acids and stable isotopes data. For the study, two types of samples were analyzed, subcutaneous fat from the alive animal and from the slaughtered animal. The first type of samples requires a biopsy on the animal that needs to be done by a specialized veterinarian. This shows that NIRS technology is also useful for producers, as the quality of the ham to be sold and, therefore, its price can be determined before the production. However, this technique on live animals requires specialized personnel that increases the cost of the analysis.

Fernández Novales [57] proved that the use of portable NIR sensors could correctly classify 96% of 367 different Iberian ham samples according to its fat content and 83% according to the muscle content both previous and after the salting process. Tejerina [58] proved that these portable sensors are also suitable for salt content with a mathematical adjustment related to the absorbance of the organic matter of the sample, as inorganic compounds do not absorb NIR.

NIRS portable sensors are an emerging tool in the detection of possible frauds in this field, such as the commercialization of crossbreed hams or counterfeiting.

Another example of an application of NIRS sensors in Spain is on the field of wine vinegars. In Andalucía, there are three PDO related to wine vinegars, which are: *"Vinagre de Jerez"*, in Cádiz; *"Vinagre del Condado de Huelva"*, in Huelva; and *"Vinagre de Montilla-Moriles"*, in Córdoba [59]. The latter has been found to be the target of many investigations. Back in 2014, de la Haba [60] confirmed that NIRS technology could ensure if a vinegar meets the legal requirements for PDO *"Vinagre de Montilla-Moriles"*. Márquez Ortega [61] also showed the power of this sensors to determine quality and authentication for this PDO. Moreover, Ríos-Reina [62] determined that NIRS technology could classify the samples in its corresponding PDO (*Condado de Huelva*, *Montilla-Moriles*, *Jerez* and no PDO) in more than 90% of the cases.

It could be therefore concludable that NIRS sensors are a good alternative for fraud detection in wine vinegars for being versatile, cheap, and noninvasive. These sensors can be incorporated in different decision levels along the chain of production [61].

Examples can also be found outside of our borders. In Italy, Gragnano pasta is a PGI from a small zone in the province of Naples. It is considered the best pasta in the world. It has been analyzed through conventional methods such as GC-MS [63]. Recent technologies have shown that NIRS technology also applies for this product, avoiding food frauds and authenticating it [64]. NIRS has also been used in Italy for correct identification of PDO hazelnut variation *"Nocciola Romana"* [65]. Other food products for which NIRS has been used can be Halloumi cheese PDO in Cyprus [66] and Serrana and Preta de Montesinho meats in Portugal [67].

As it can be seen, NIRS is a very powerful tool for food fraud detection with different characteristics: being noninvasive, fast, and applicable in different stages of the production chain.

5.- METHODS

5.1.- OIL SAMPLES

Seven different types of oils were employed in this work. On the one hand, three varieties of olive oil were used: extra-virgin olive oil (EVOO), virgin olive oil (VOO) and olive oil (OO). On the other hand, four different seed oil adulterants were used for sample preparation: sunflower oil, flax oil, sesame oil and mix of seeds A total of 160 samples were analyzed, all provided by GEAB investigation group of the University of Oviedo having carried a similar experiment beforehand [15]. As it is shown in Table 5.1, samples were prepared adding different seed oils. The binary samples preparation procedure consisted in mixing pure olive oils with an adulterant oil changing the proportions between them. The 48 binary samples had a percentage of adulterant oil ranging from 2 to 30%. For the 115 ternary samples, pure olive oils were mixed with two oils of the four possible adulterants, varying the proportion of adulterated oils between 3 and 16%.

Table 5.1- Percentage distribution of the binary and ternary oil samples

Binary samples		Ternary samples		
% Olive oil	%	% Olive oil	%	%
98	2	90	5	5
90	10	80	10	10
80	20	70	15	15
70	30	88	9	3
		76	8	16
		92	5	3
		86	11	3
		82	10	8

5.2.- NIR SPECTRAL COLLECTION

The oil samples were analyzed with an assembly simulating an online process. Figure 5.1 depicts the different parts of the experiment.

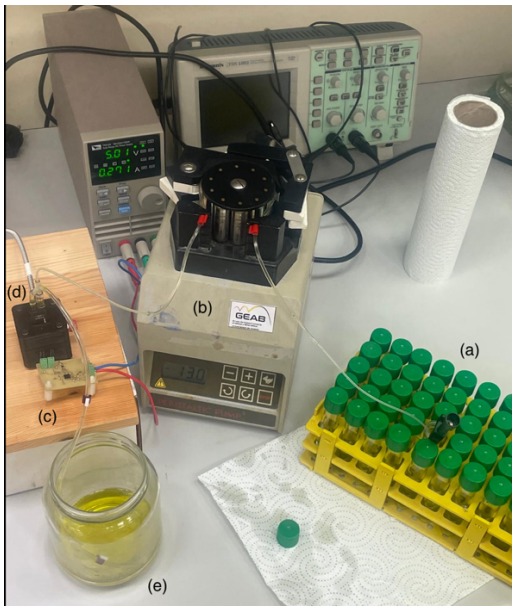

Figure 5.1- Assembly for spectral collection with: (a) Sample rack, (b) hydraulic pump, (c) lamp, (d) cuvette, (e) waste flask.

Each of the samples was pumped from the sample container to the flow injection analysis cuvette by using a peristaltic pump. Once the cuvette is filled and the oil is in the exit of the flexible pump tube, the spectrum is collected. All samples were measured without any pretreatment, at room temperature and in triplicate. Between two different oil samples, a small flow of air was let by in the circuit. If the measurements were taken on different days, the circuit was cleaned with methanol. Measurements were carried out in the wavelength range from 901 to 1600 nm with 195 different points.

Before starting to collect sample spectra, the experimental conditions were optimized. The peristaltic pump was calibrated to obtain the relationship between the angular velocity provided by the pump and the speed of the fluid in the system. The exit speed of the oil was measured by collecting a known volume with a graduated cylinder and dividing it for the time taken to reach it.

After that, the following step was to optimize the speed of the flow and the number of scans to be averaged by comparing the repeatability of the collected spectra.

The higher the speed of the pump, the faster the cuvette is filled, and the faster samples can be measured. However, excessively high speeds can induce turbulences in the cuvette and affect the repeatability of spectra results. Another parameter that must be optimized because it affects the quality of collected spectra is the number of sub-scans to be averaged to obtain the final representative spectrum of each olive oil sample. This parameter is a number which is implemented in the software of the instrument. A low average of sub-scans results in a faster scan time, but the data may lack precision.

To select the best experimental conditions to collect all the spectra, the root mean square error (RMS) statistic was used. This statistic compares pairs of spectra, and its value can be associated to the reproducibility and repeatability of the model. If the RMS is low, the difference between two different spectra is low. This means that the spectra are similar between them [15], [68]. RMS is calculated as shown in equation 5.1.

$$\mathrm{RMS}=\sqrt{\frac{\sum_i^j (y_a\text{-}y_b)}{n}}\times 10^6 \qquad (5.1)$$

Where n is the number of wavelengths, y_a the absorbance for the spectrum a at each wavelength, from wavelength i to wavelength j and y_b is the absorbance for the spectrum b at each wavelength (from i to j).

5.3.- SPECTRAL DATA PROCESSING

The spectra obtained with the device were transformed into a data matrix with two variables: the wavelength and the absorbance data. To obtain this matrix, the text files provided by the software were transformed into a spreadsheet using *Microsoft Excel 2021*. The arithmetic mean was computed between the three spectra collected.

The chemometric approaches were developed using *Unscrambler X* software (The Unscrambler X, CAMO Analytics AS, Oslo, Norway). An example of the program screen is displayed in Figure 5.2.

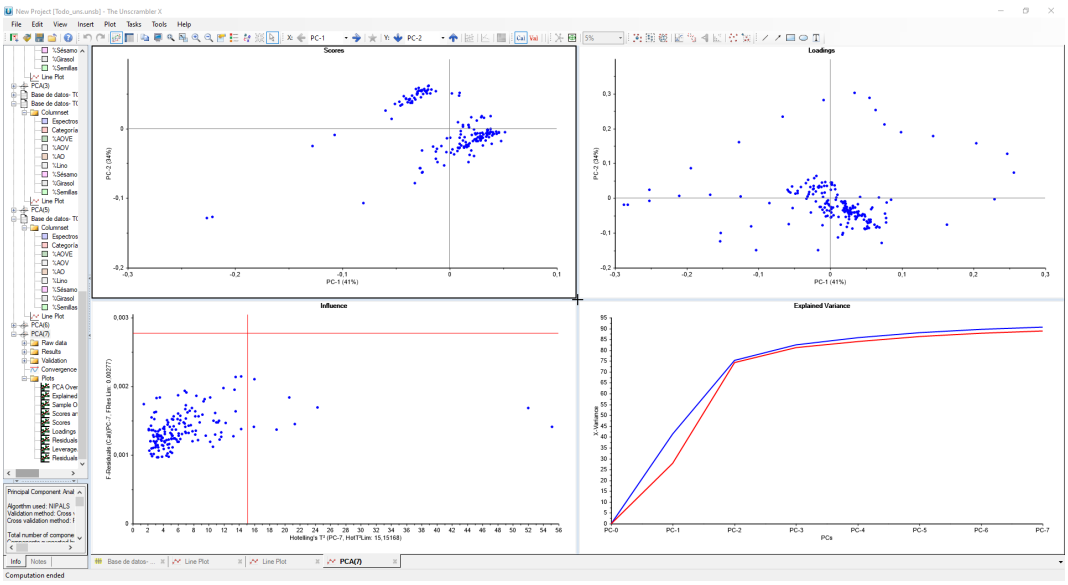

Figure 5.2- Unscrambler X interface

The strategy to follow was to define three categorical variables: olive oils, seed oils and adulterated samples. This approach allows the identification of presence of seed oil adulteration on olive oil.

The chemometric approach was based on applying the methods explained in section 4.2. Standard normal variate (SNV) was applied to minimize light scattering and Savitzky-Golay derivatives (SG) were used to reduce noise and eliminate the baseline. And after that, Principal Component Analysis (PCA) was used for the detection and classification of the adulterations. All the developed models were optimized using a random cross-validation method included in chemometric software *Unscramble X*. Table 5.2 shows all the pre-treatments combinations carried out to develop PCA classification models.

Table 5.2- Chemometric approaches assayed.

Treatment number	Derivative	Derivative order	Polynomial order of derivative	Left interval size	Right interval size	SNV
1	SG	1	2	4	4	No
2	SG	2	2	4	4	No
3	SG	1	2	4	4	After
4	SG	2	2	4	4	After
5	SG	1	2	4	4	Before
6	SG	1	2	4	4	Before

The nomenclature to be adopted from this point for the tested pre-treatments will be as follows: SG A-B-C-D. A is the Savitzky-Golay derivative order (the tested values were 1 or 2), B is the polynomial order of derivative; and C and D are the number of smoothing points taken from the left and right, respectively.

6.- INSTRUMENTATION

6.1.- CUVETTE

The cuvette is Hellma 1767008510-40 flow-through cell (Hellma Analytics, Müllheim, Germany). It can be seen in Figure 6.1.

Figure 6.1- Hellma flow-through cell. From [69].

It is a quartz-glass high performance cell with a wavelength range from 200 to 2500 nm. It has a 10 mm optical path length and a volume of 390 μL. Its outside dimensions are 35x12.5x12.5 mm (height x width x depth) [69].

6.2.- SPECTROPHOTOMETER

The spectrophotometer used is DLP NIRscan Nano spectrometer (EVM, Texas Instruments Incorporated, Dallas, TX, USA). It is connected to the receptacle holder showed in Figure 5.1 with a fiber optic wire. The device is also connected to a PC with a USB wire so that the software provided by the manufacturer can be used. The dimensions of the mini device can be seen in Figure 6.2.

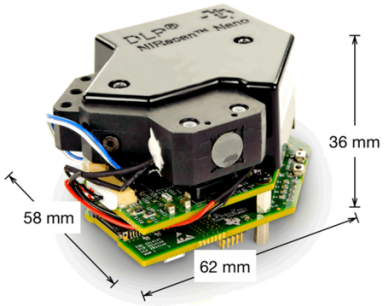

Figure 6.2- Dimensions of the DLP NIRscan Nano spectrometer.

The DLP (digital light processing) based spectrometer does not use a typical linear array detector, but a digital micromirror microdevice (DMD). First, each of the wavelengths is separated and the data are collected point by point. The single points are then collected in the DMD and rearranged to form the spectra. Figure 6.3 shows the differences between a conventional array detector and the use of a DMD.

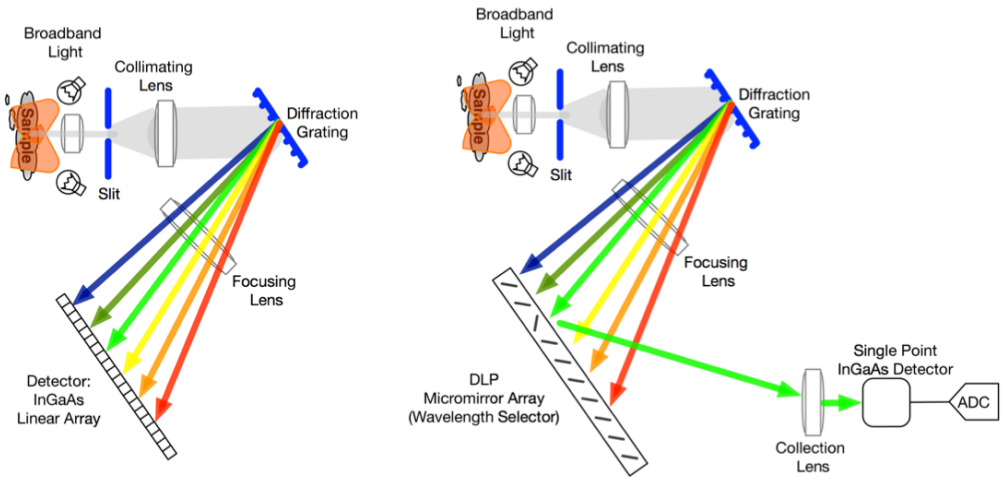

Figure 6.3- Difference between linear array and DLP-based spectrometer.

DLP technology is based on measuring the spectrum of light on an incident port. The incident beam of light is dispersed in different wavelengths and then reimaged in the DMD. This allows higher performance than a conventional linear array. The DMD also corrects the distortion of each individual system.

The main advantage of this system is that it allows fast, flexible, and programmable spectral filters. Figure 6.4 shows the optical parts of the spectrometer [DLP NIRscan™ Nano EVM User's Guide. Texas Instruments].

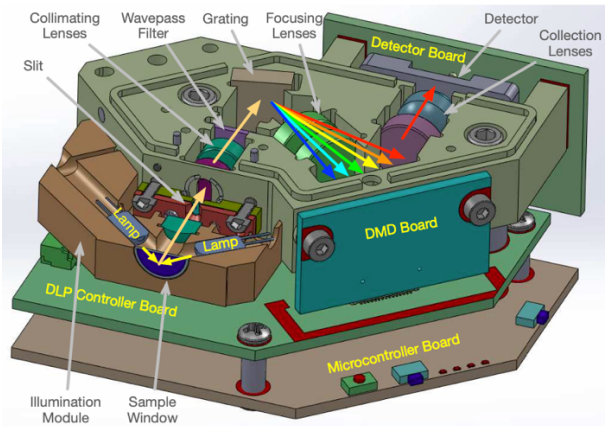

Figure 6.4- DLP NIRscan Nano Optical Engine.

The lamp used was a tungsten lamp and measures were taken in transmittance mode. All information and figures of this subsection have been taken from the manufacturer user guide [70] and the design considerations [71].

6.3.- NIR SOFTWARE

One of the best advantages that the spectrometer has is that it provides a specialized software for the collection of the spectra. Figure 6.5 shows the screen displays including the instrumental parameters to measure NIR spectra. All spectral data were collected as the absorbance at each wavelength.

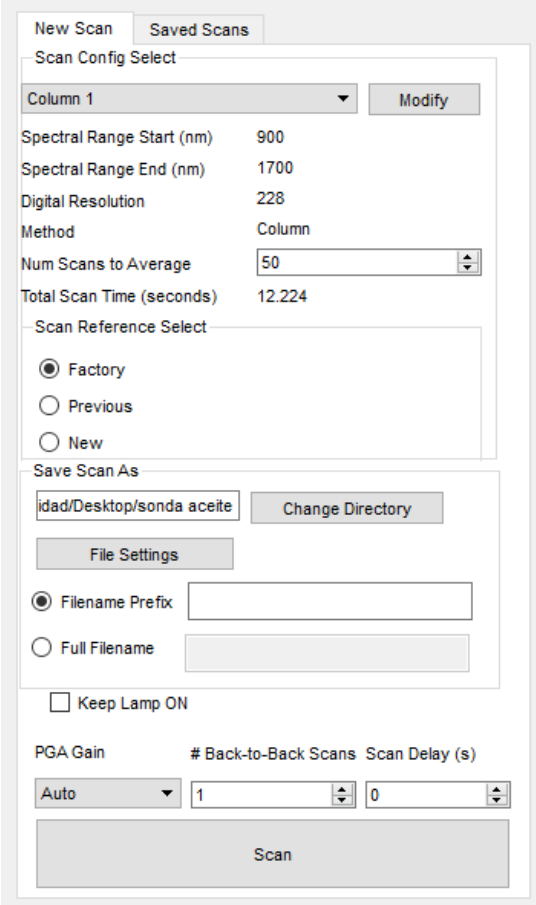

Figure 6.5- Scan parameters screen for DLP NIRscan Nano software.

Among the most important parameters (see Figure 6.5) to be checked in the software are the measuring method, the wavelength range, the number of scans to average and the total scan time. The two different modes of scan are Column 1 and Hadamard. Column 1 method was used, as previous research showed it was the most appropriate one for solving the analytical problem [15]. The wavelength range was set between 900 and 1700 nm. However, from 1500 to 1700 nm showed errors (noisy spectra) during the spectral collection and were avoided having considered that it was not a region of interest.

7.- RESULTS

7.1.- PERISTALTIC PUMP CALIBRATION

As explained above, the peristaltic pump was calibrated to ensure that the angular speed was lineally related to the oil flow. To measure the flow, volume and time were measured with a graduated cylinder. The linear correlation was strong between flow and pump angular speed attending to the R^2 statistic parameter.

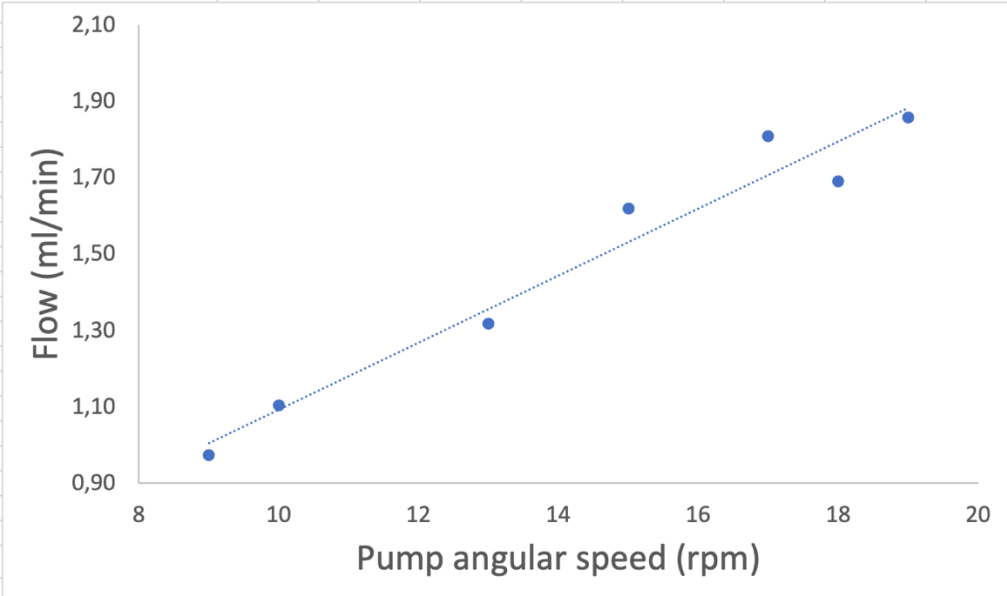

Figure 7.1- Peristaltic pump calibration.

7.2.- OPTIMIZATION OF THE PARAMETERS

To ensure the quality of the spectral collection, two parameters were optimized: flow speed and the number of scans to be averaged. As detailed in section 5.2, RMS was used as statistical parameter to detect the best reproducibility of the model. Olive oil was pumped through the experimental system, five repeated measurements were taken for each condition and RMS values were calculated. Results are shown in Figure 7.2.

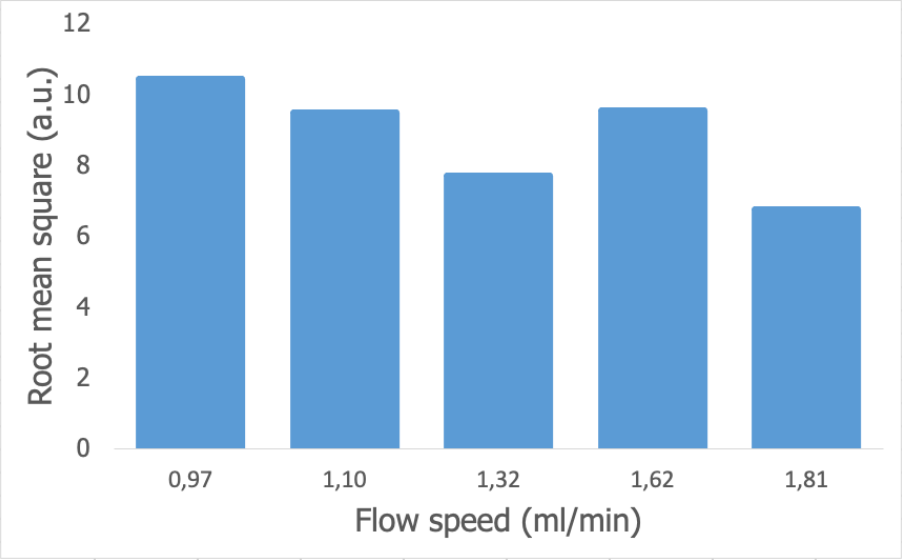

Figure 7.2- Flow speed optimization.

The lowest RMS was found to appear at 17 rpm (1.81 ml/min). However, it was found when oil was put through the circuit, due to its lubricant character, an overpressure caused flexible pump tube to slip and come off due to the internal size of the cuvette. This made necessary to change the speed to 13 rpm, as it was the second lowest RMS value. The spectra already collected at 17 rpm (1.81 ml/min) were discarded and all measurements were taken at a pump speed of 13 rpm (1.32 ml/min).

The second step was to select the number of sub-scans to be averaged to obtain a final spectrum. It was firstly evaluated at 17 rpm, which is shown in Figure 7.3. As this speed was proven not to be suitable for the experiment, it was then tested at 13 rpm, shown in Figure 7.4.

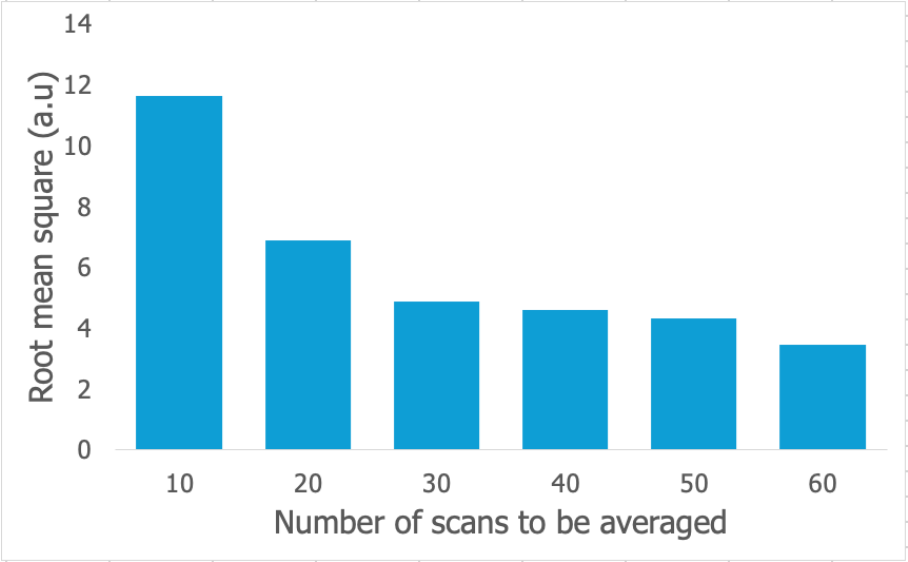

Figure 7.3- Number of scans to be averaged at 17 rpm.

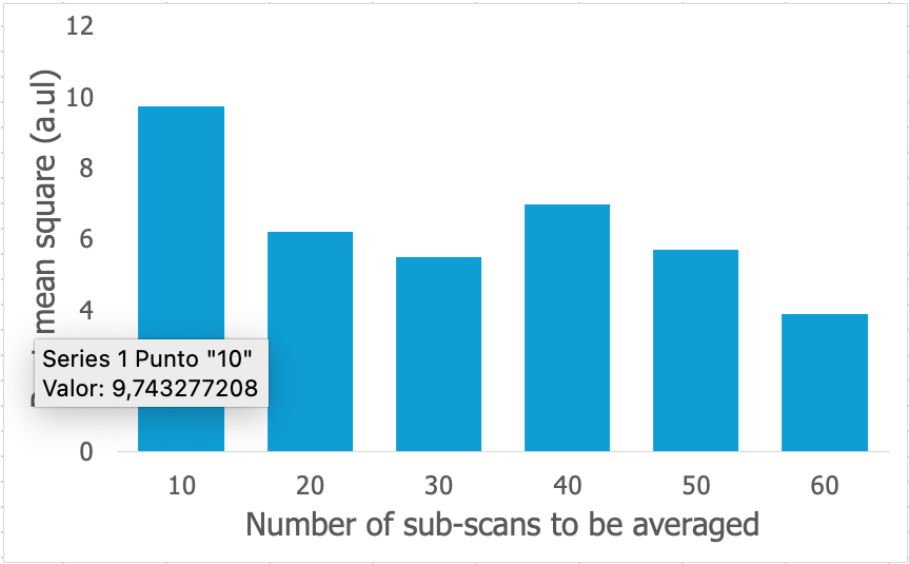

Figure 7.4- Number of scans to be averaged at 13 rpm.

Figure 7.3 shows lower RMS values than Figure 7.4, as a speed of 17 rpm had a lower RMS value than 13 rpm. As can be seen in Figure 7.4, similar RMS results were obtained for 30, 50 and 60 scans to average. During the assay, we observed that 30 scans to be averaged gave problems every few spectra and some of them needed to be repeated. This did not occur when averaging 50 and 60 (which show low RMS value). Finally, the value 50 was chosen to minimize the analysis time. The measuring time for each spectrum at 50 scans to be averaged was 12 seconds.

Table 7.1 summarizes the final conditions to be used during the experiment as a summary for this subchapter.

Table 7.1- Final operating conditions.

Optimized conditions	
Peristaltic pump angular speed (rpm)	13
Flow speed (ml/min)	1.32
Number of scans to be averaged	50

7.3.- CHEMOMETRIC TREATMENTS

After setting optimal conditions in the previous section, the spectra of all samples were collected in triplicate. Collected spectra are shown in Figure 7.5 (a), while Figure 7.5 (b) shows the spectra of only pure olive oils and pure seed oils. The first PCA was tested using raw spectra (Figure 7.6).

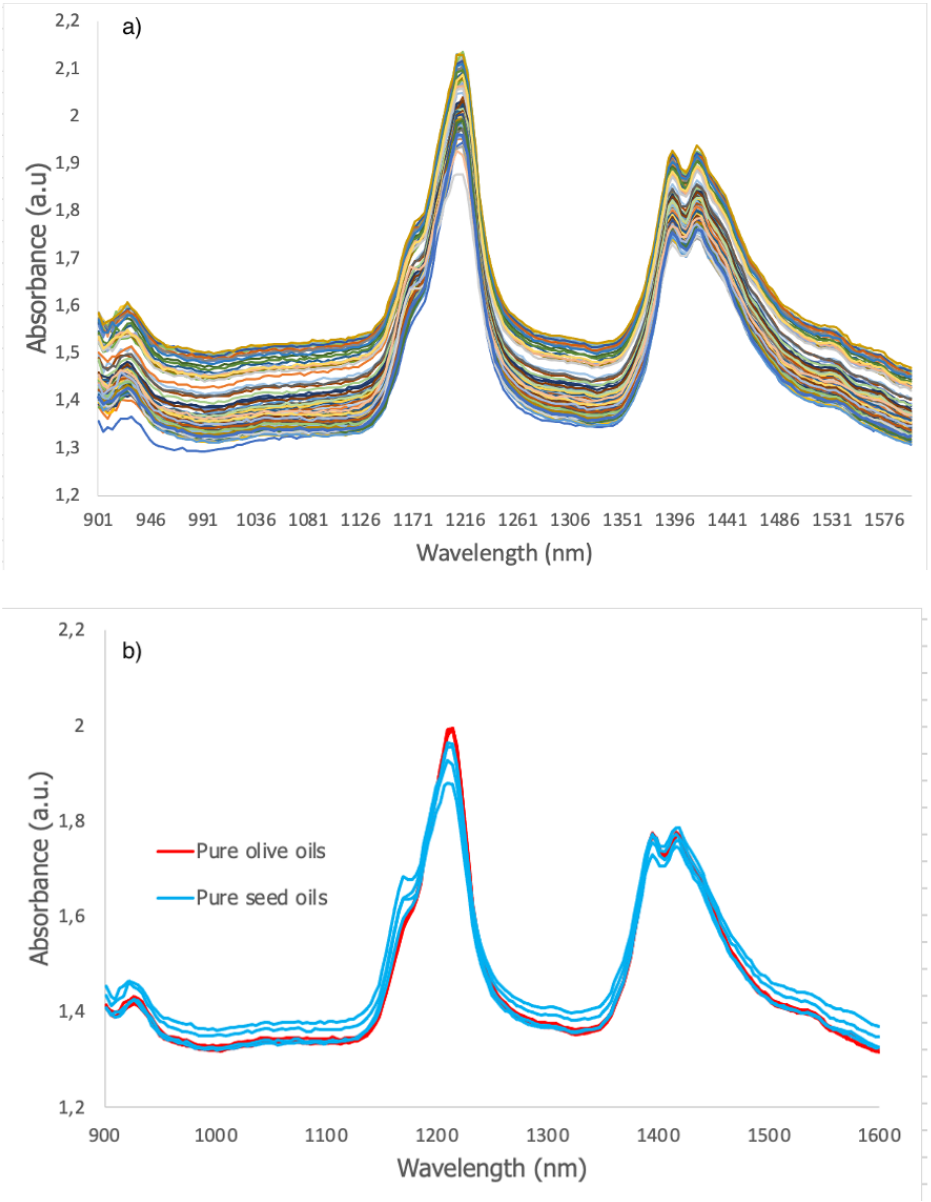

Figure 7.5- Spectral plots of the raw data. a) All data. b) Pure olive oils and pure seed oils

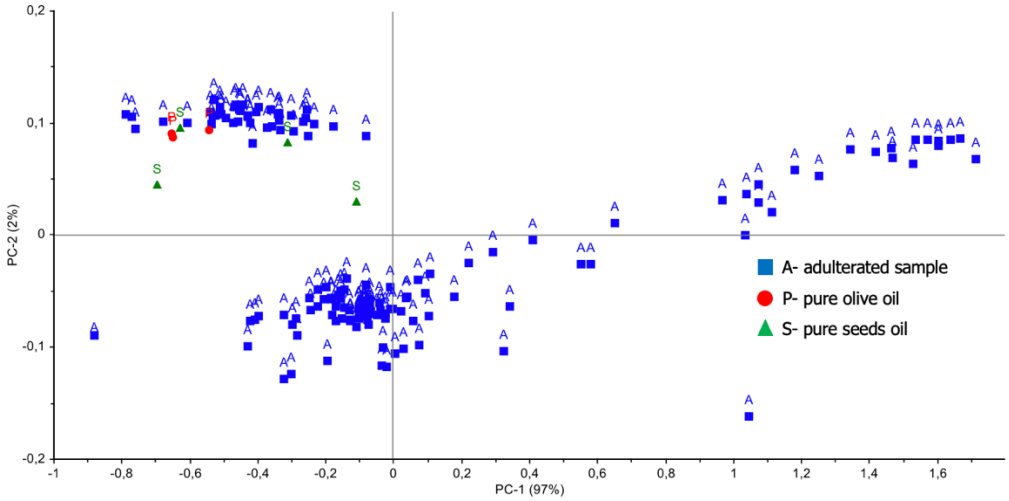

Figure 7.6- PCA plot of the raw data.

As can be seen in Figure 7.6, the green triangle markers correspond to pure seed oils (S), the blue squares to adulterated samples (A) and the red dots to pure olive oils (P). This legend will be used in all PCA plots during the development of this chapter. PCA using raw spectra (see Figure 7.6) showed unsatisfactory results, as pure olive oil samples cannot be differentiated from adulterated samples. Therefore, the different math treatments are applied to extract spectral relevant information and to obtain PCA plots where pure olive oil samples can be visually distinguished from all other samples. Prior to the PCA classification the effect of the math pretreatment on spectra data are shown as figures exported from Unscramble X software (see Figure 7.7 and Figure 7.8).

Figure 7.7 and Figure 7.8 show the three spectral plots without applying SNV and applying SNV prior and after 1st and 2nd order Savitzky-Golay derivative. As can be seen, these spectral plots do not give relevant information for the distinction of adulterated and non-adulterated samples. However, useful data can be extracted from the graphs. For example, the baseline effect observed on raw data (Figure 7.5) is corrected with the use of both derivatives. Also, all spectra are centered on the same line, noise is reduced, and all spectra are smoothed. Therefore, it can be affirmed that the spectral plots shown in both figures (Figure

7.7 and Figure 7.8) graphically demonstrate the effect of the applied math pre-treatments (section 4.2).

The main visual difference that can be extracted from these graphs is related to the SNV treatment. Figure 7.7 shows that applying the SNV treatment results in more accentuated peaks (Figure 7.7 (b) and Figure 7.7 (c)) than not applying it, as in Figure 7.7 (a). This effect is not observed on Figure 7.8. After applying these math pre-treatments on spectra (160 spectra) PCA model was performed.

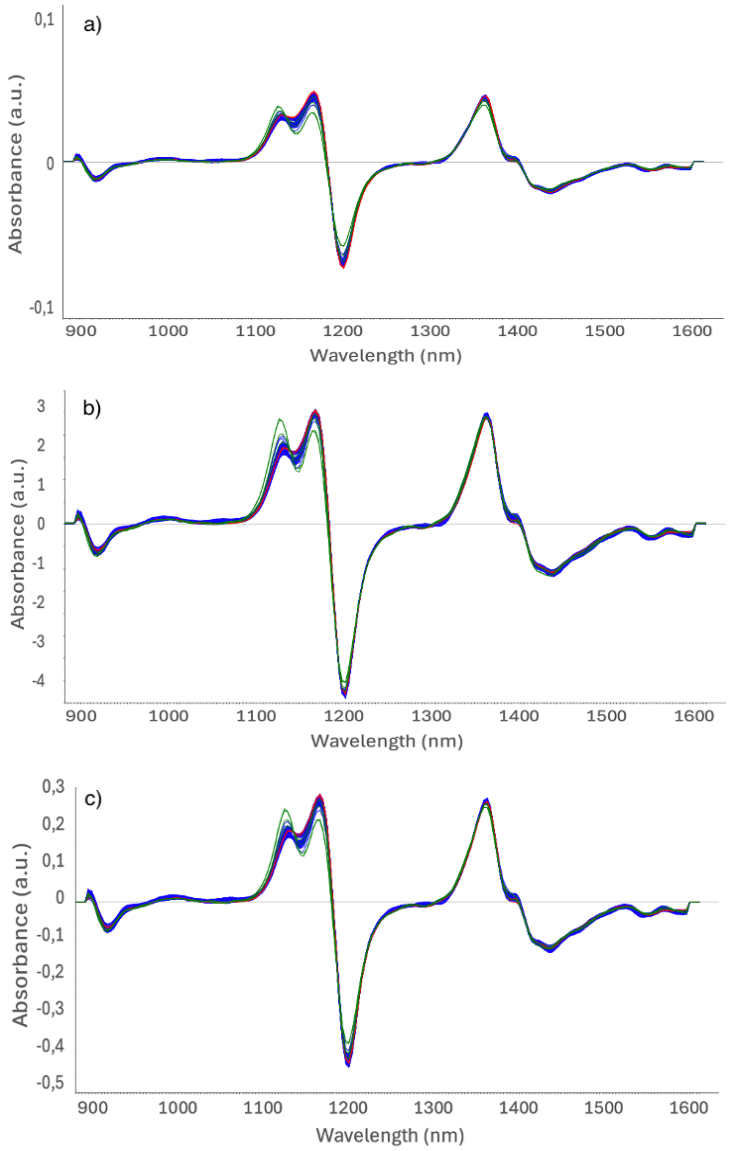

Figure 7.7- Spectral plots after chemometric treatments with first order derivatives. a) SG-1-2-4-4. b) SG-1-2-4-4 + SNV. c) SNV + SG-1-2-4-4.

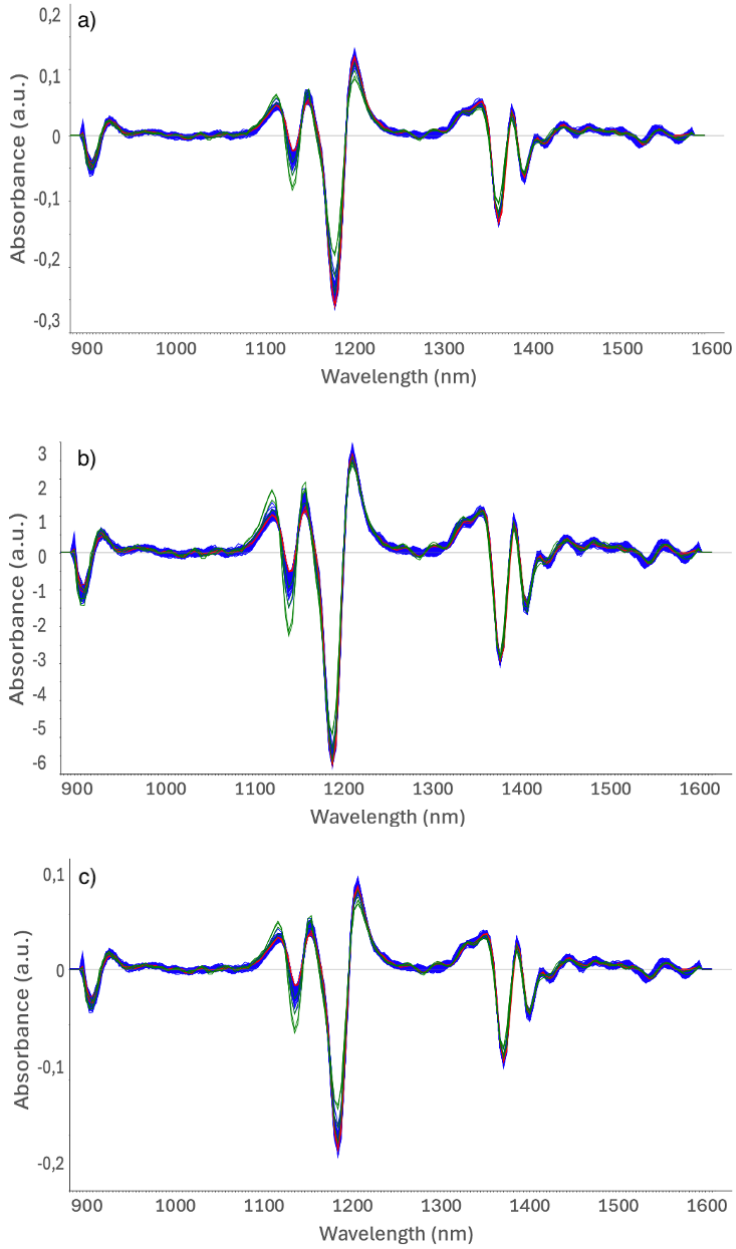

Figure 7.8- Spectral plots after chemometric treatments with second order derivatives. a) SG-2-2-4-4. b) SG-2-2-4-4 + SNV. c) SNV + SG-2-2-4-4.

Before PCA was carried out, the spectra of olive oils and seed oils were analyzed separately (without adulterated samples) to look for any spectral difference. As an example, only one of the treatments was chosen, SG 1-2-4-4. The result can be seen in Figure 7.9.

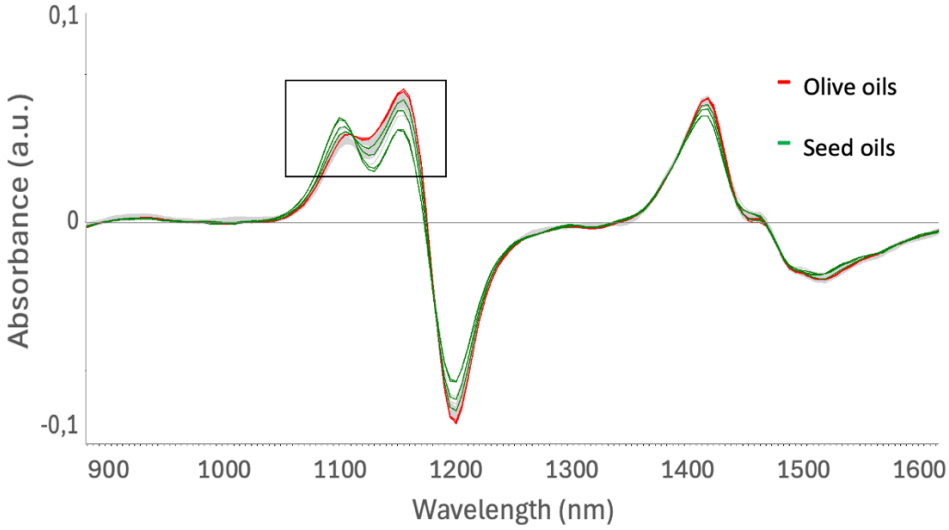

Figure 7.9- Spectral differences between pure olive oils and pure seed oils.

The main visual difference observed in Figure 7.9 appears at approximately 1160 nm (see the black box of the figure). Seed oils present two more distinguished peaks with a valley between them, whereas olive oil shows only one peak with a shoulder at that wavelength. This peak around 1205 nm corresponds to the second overtone vibration of the C-H (CH_2) link [72]. This band has been established by previous authors as a characteristic band in olive oil samples [73].

PCA analysis was subsequently performed. The result of the first PCA is shown in Figure 7.10, which corresponds for the first mathematical pre-treatment, that is SG-1-2-2-4. Figure 7.10 shows clear differentiation between the set of pure olive oils and the rest of the samples. In PCA score plots, the closer the points are, the more similarities they have. Figure 7.10 shows three groups, one cluster of red dots corresponding to olive oils and two big clouds of adulterated samples.

The points of the spectra are transformed into a group of principal components, that explain the variability between them. The components chosen for the plots are always the two that represent the variability between spectrum in the most accurate way. For example, in Figure 7.10, the principal components represented

in the x-axis and y-axis are the two that better represent the variability between the samples. The percentage refers to the fraction of the total variability represented by the principal component. Principal component 1 represents 48% of the total variability and principal component 2 represents 32%. As can be seen in Figure 7.10 olive oils (P) appear separately of the rest of adulterated oils or seed oils after applying SG first derivative.

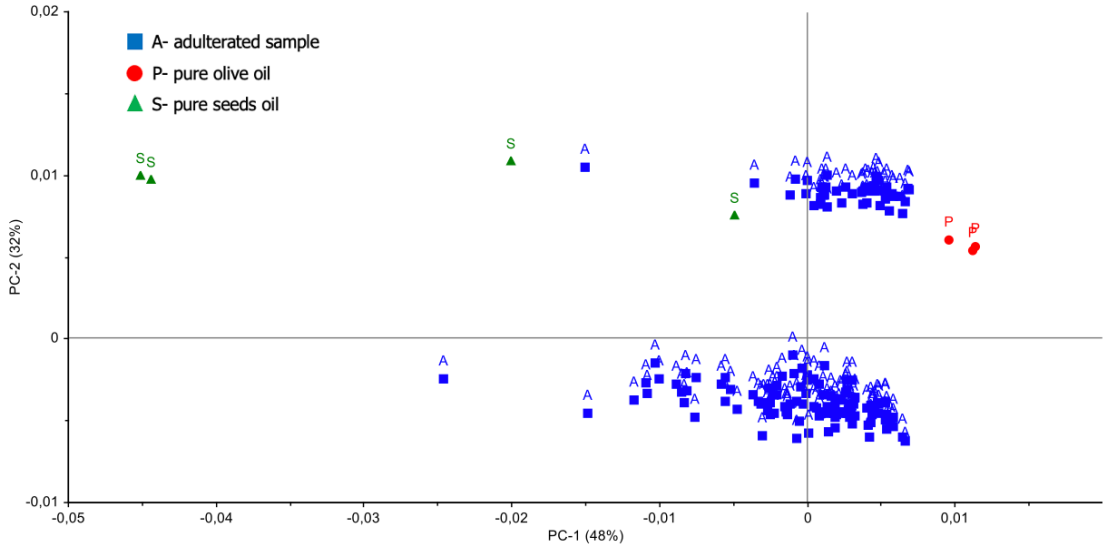

Figure 7.10- PCA plot with math pre-treatment SG-1-2-4-4.

The following step was analyzing the effect of applying the SNV pretreatment after and before the derivative (SG-1-2-2-4 + SNV and SNV + SG-1-2-2-4). Figure 7.11 shows both PCA plots. Again, the three points corresponding to pure olive oil samples can be easily distinguished. No significant differences are found between applying the SNV treatment prior to the derivative and applying it after the derivative.

The most eye-catching point is the one corresponding to sesame oil, which is highlighted in Figure 7.11 (a) and Figure 7.11 (b) with a black circle. This point appears to have a behavior that is halfway between pure olive oils and the rest of pure seed oils. However, the fatty acids profile of sesame oil shows similarities

with that of the olive oil [74]. Table 7.2 shows the similarities between the two oils.

Table 7.2- Similarities between sesame oil and olive oil. Adapted from [74].

Fatty acids	% in sesame oil	% in olive oil
Saturated fatty acids	10.5-18	14
C14:0 (myristic acid)	0	0
C16:0 (palmitic acid)	7-12	10.8
C18:0 (stearic acid)	2.5-6	2.8

In general, all three pretreatments involving first order derivatives show very strong results, as all pure olive oils can be well distinguished to the rest of the samples and form a small group between them with all three points very close to each other. Moreover, the three pretreatments show results very similar between them, attending at the distribution of the points. No differences can be seen between not applying the SNV pretreatment, applying it after the derivative and applying it before.

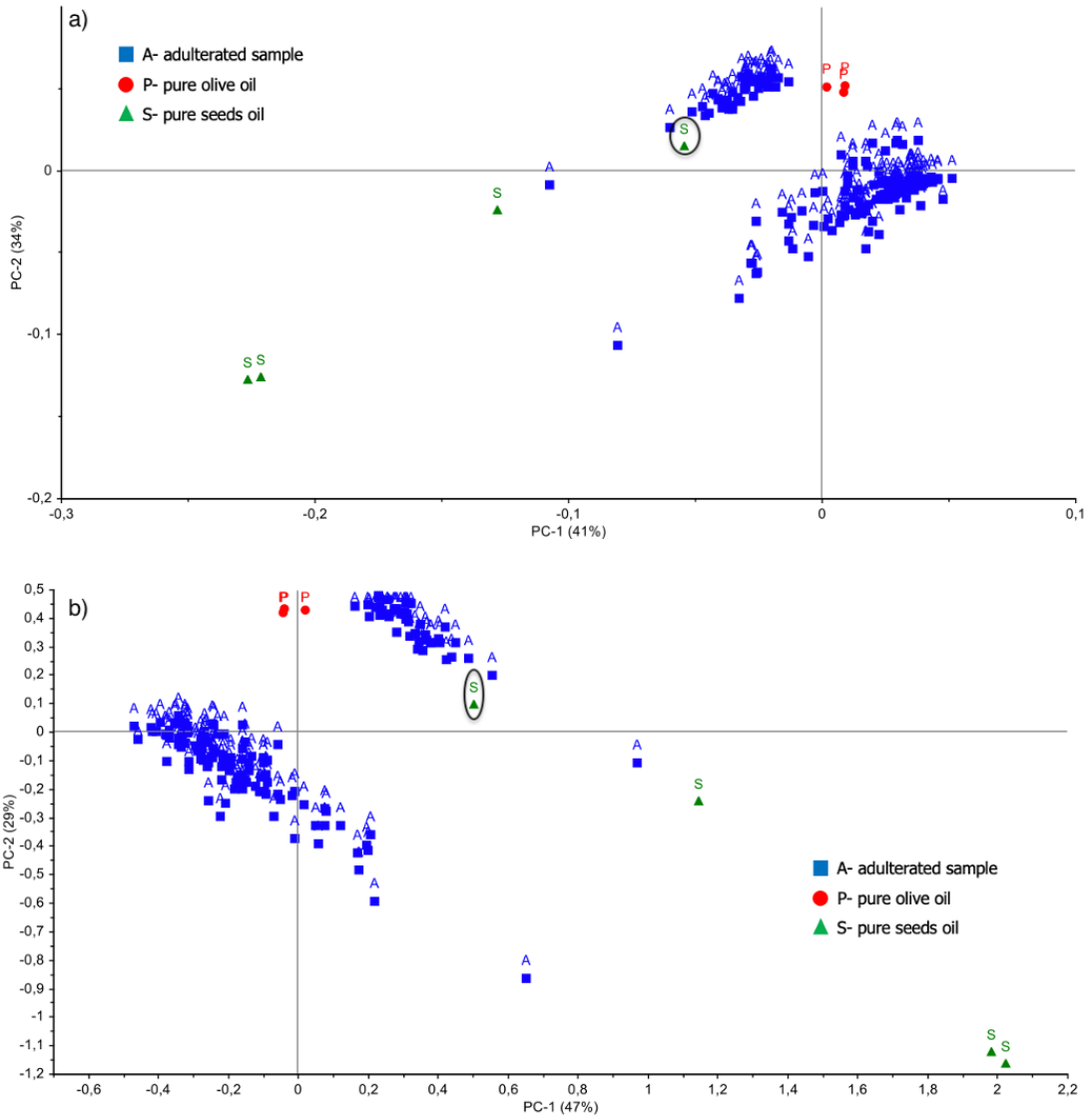

Figure 7.11- PCA plot with math pre-treatments (a) SG-1-2-4-4 + SNV. (b) SNV + SG-1-2-4-4.

Once the three pretreatments with first order derivatives were analyzed, PCA was performed on pretreatments with second order derivatives. Figure 7.12 shows the first pretreatment (SG-2-2-4-4). The three points corresponding to olive oils can be again visually distinguished. However, they do not form a cluster as seen in Figure 7.10 and Figure 7.11. There is a point that, although it still lies separately to the adulterated samples, is also separated to the other two olive oils point. This point, contained in a black circle, corresponds to olive oil (OO)

while the other two that lie together correspond to extra virgin olive oil (EVOO) and virgin olive oil (VOO). This poorer result can be explained with the PC percentage, as both principal components only represent 50% of the variability, (32% and 18%)

There are also two points that correspond to pure seed oils and have similar behavior, which are highlighted with a red box in Figure 7.12. This was also seen in Figure 7.10 and Figure 7.11. These two points correspond to sunflower oil and mixed seed oils.

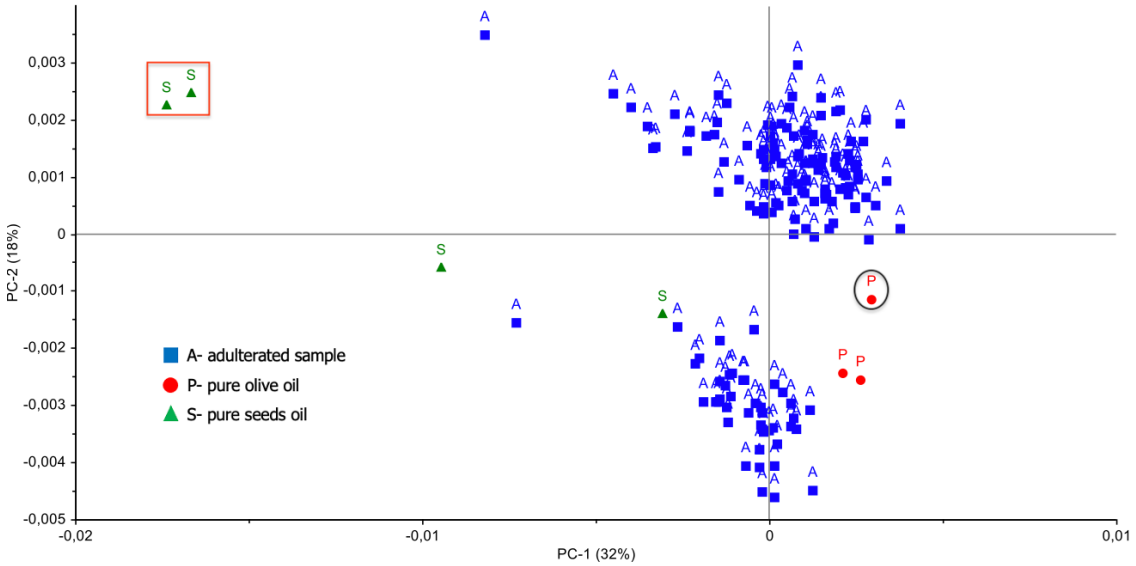

Figure 7.12- PCA plot with math pretreatment SG-2-2-4-4.

The effect of the SNV pretreatment was also tested in second order derivative pretreatments. The result is shown in Figure 7.13, where figure 7.13 (a) represents applying it after the derivative and figure 7.13 (b) represents applying it before the derivative.

In this case, there are differences between both pretreatments. Figure 7.13 (a) shows similarities with Figure 7.12, where the points can be distinguished, but

OO is separated from EVOO and VOO. However, figure 7.13 (b) shows very bad results. In this figure, some points corresponding to adulterated samples form one small cloud while the rest are dispersed along the graph. The points corresponding to pure olive oils are mixed with this cloud of points, not making possible the identification between them. No relevant information can be extracted from figure 7.13 (b).

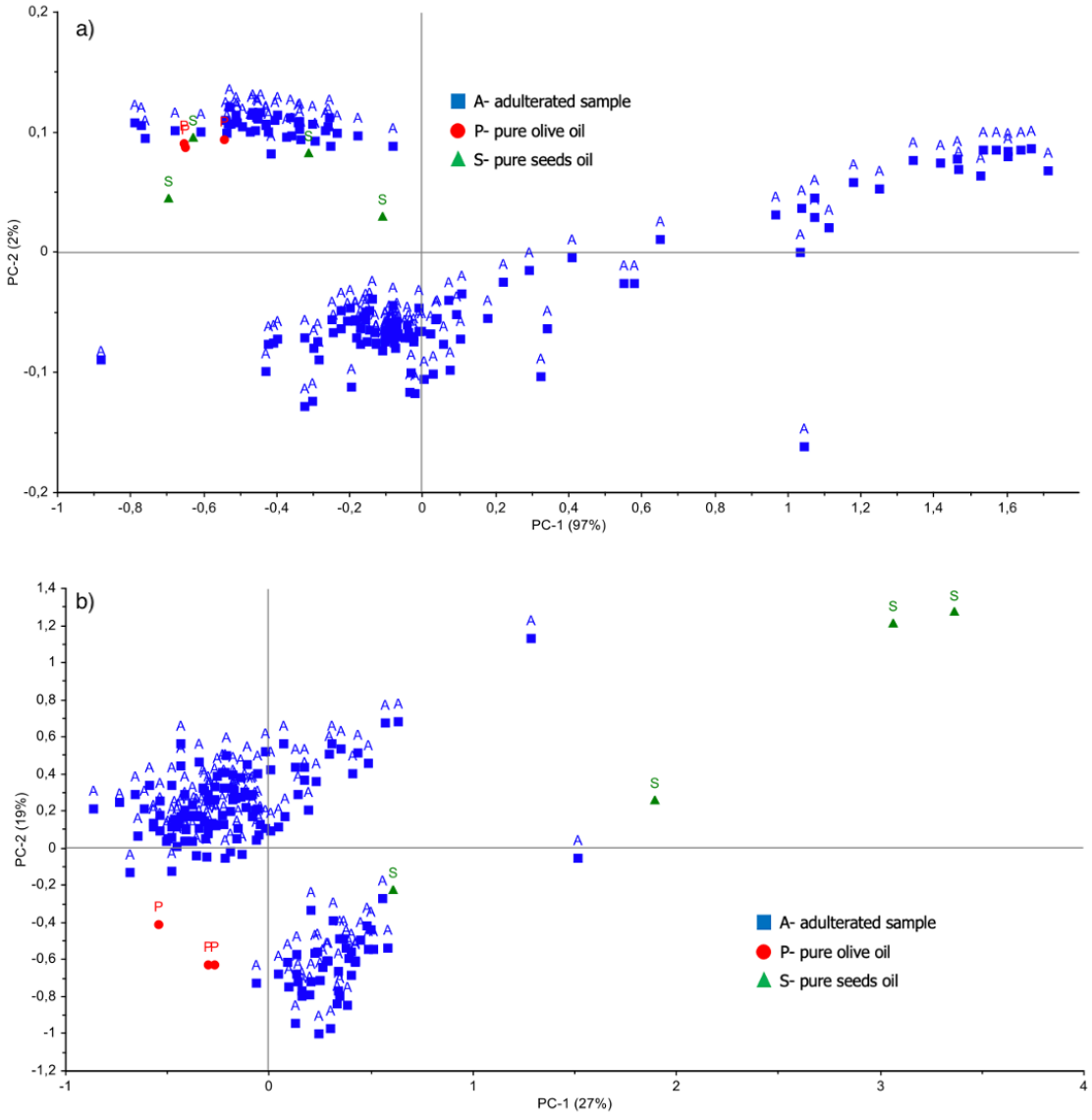

Figure 7.13- PCA plot with math pre-treatments (a) SG-2-2-4-4 + SNV. (b) SNV + SG-2-2-4-4.

Figure 7.11 (1st derivative Savitzky Golay combined with SNV) shows that no difference is present between applying the SNV treatment before and after the derivative if the pretreatment involves first order derivatives. However, it does not happen when the treatments involve second order derivatives, as Figure 7.13 (a) and (b) show completely different results.

Table 7.3 shows a summary of the quality of the chemometric models developed.

Table 7.3- Summary of chemometric models developed.

Number	Derivative	SNV	Result
1	SG-1-2-4-4	No	Excellent
2	SG-1-2-4-4	After	Excellent
3	SG-1-2-4-4	Before	Excellent
4	SG-2-2-4-4	No	Good
5	SG-2-2-4-4	After	Good
6	SG-2-2-4-4	Before	Bad

Recent research on olive oil fraud detection with NIRS [15] showed that samples containing the mixture of the three seed oils could cause inconveniences when performing PCA. All chemometric models were tested after eliminating this kind of samples from the data matrix.

As the poorest results previously analyzed were the ones involving second order derivatives, their PCA plots are shown in Figure 7.14. The treatment with no SNV applied, which is Figure 7.14 (a), shows no difference in results to that without eliminating samples with the mixed seeds oil Figure 7.12. The same happened with Figure 7.14 (b) and Figure 7.13 (a), with the same results if the SNV was applied after the derivative for both the data matrix with mixed seeds oil and without it.

However, there is a difference between Figure 7.14 (c) and Figure 7.13 (b). When applying the SNV before the derivative, very poor results were obtained when including mixed seed oils. This was corrected in Figure 7.14 (c), as the results eliminating this type of samples are strong. The mixed seed oil was causing the pretreatment to fail. When eliminating this type of samples, no difference is found between applying SNV before and after the derivative.

Annex I shows the spectral plots when eliminating the mixed seed oil. Annex II shows their PCA score plots with first order SG derivatives pretreatments. Both show similar results to those when those samples are not eliminated from the data matrix.

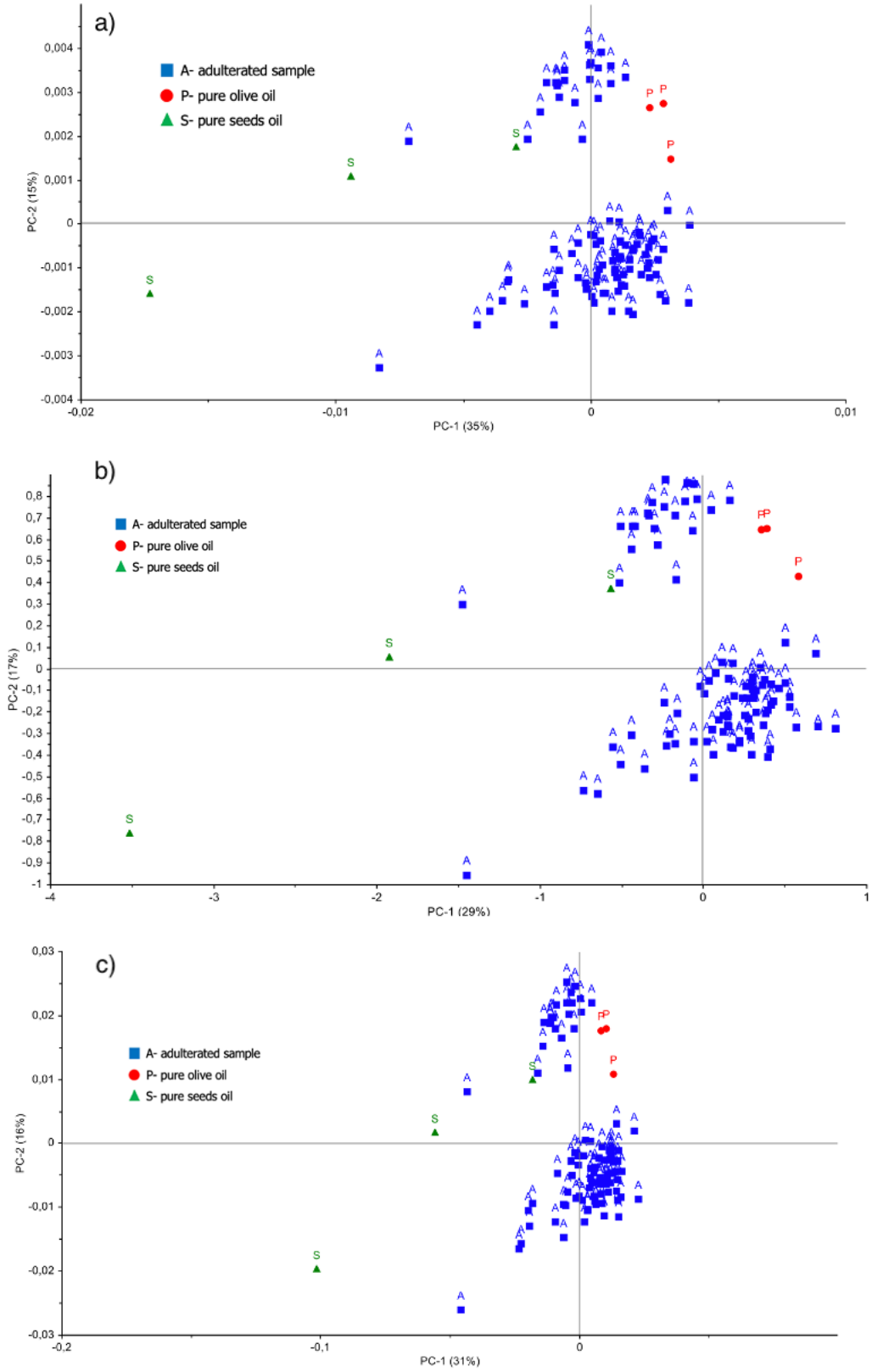

Figure 7.14- PCA plots of math pretreatments with second order derivatives (with no seeds oil). a) SG-2-2-4-4. b) SG-2-2-4-4 + SNV. c) SNV + SG-2-2-4-4.

To summarize, very good results are obtained when applying pretreatments with first order derivatives, as there is a small cluster that contains all olive oil samples and is differentiated from the rest of samples. With this type of pretreatments, no difference is seen between applying the scatter correction (SNV) or not.

Regarding second order derivative pretreatments (Figure 7.12 and Figure 7.13), the results are not satisfactory, but still allow the differentiation of pure olive oils. In this case, the three points do not form a cluster, as OO is separated from EVOO and VOO. The treatment when applying second order derivatives and SNV before it showed bad results. However, research showed that the mixed seeds oil could cause these problems. When eliminating spectra of samples containing it, the pretreatment was corrected and showed the same results as the one with no SNV and the one with SNV applied after the derivative.

8.- CONCLUSIONS

NIRS technology is one of the most used techniques in non-invasive food analysis. It offers many possibilities, being a very flexible analytical method that has many different applications in the sector. Its main features are being fast, non-invasive and that it can be implemented on an online process. Therefore, it is a suitable technology for the analysis and detection of food fraud in olive oils.

The collection of NIRS spectra is a very simple and fast technique. However, the information obtained from raw data is not useful and require applying different math pretreatments and chemometric models to interpret these results.

Six math pretreatments were developed and evaluated to detect frauds in olive oil. Three of them used first order Savitzky Golay derivatives: one only applying the derivative, one applying the SNV pretreatment before the derivative and one applying the SNV after it. The other three were the same (no SNV, SNV before and SNV after) but with second order Savitzky Golay derivatives.

Principal Component Analysis was carried out as a chemometric strategy to establish spectral differences between olive oils and adulterated olive oils. The first three models, with first order derivative pretreatments, offered satisfactory results. The points from pure olive oil samples formed a small group that could be differentiated from the rest. With second order derivatives, worse results were obtained. However, the three points were still clearly distinguished from the rest in two of the three treatments. The one applying the second order derivative and after the SNV did not offer good results.

After eliminating samples containing mixed seeds oil, models including second order derivatives showed good results, very similar between them.

Therefore, it can be concluded that the best results are obtained with models including pretreatments with first order derivatives, and that no effect is seen when applying the SNV pretreatment, as light scatter is not relevant in this study.

The models presented can be useful to detect olive oil frauds, although there is room for improvement, as the spectral collection takes most of the time of the experiment and could be automatized. Also, more types of oils could be included in the data base to increase sample variability and improve PCA models to detect all possible adulterations.

9.- FUTURE WORKS

There are several lines of investigation to follow from this point. One is optimizing the model of this project to be implemented at industry level. This would avoid the loss of production that is now sent to analysis laboratories. Control entities could therefore perform the analyses at the production site. Moreover, the productor could also use the technology to know instantly the quality of the product to be sold and detect any anomalies during the production chain.

Another line of investigation can be analyzing other types of oil. Although sesame, flax and sunflower are very common oils used in households, the behavior of other seed oils, such as corn oil, soy oil or rapeseed oil can be used to check if the chemometric models developed during the project are applicable to them.

Finally, the quantitative analysis can be made. For example, when an official control entity detects an adulteration at a certain production site, it could use the quantification of the adulteration to establish a certain sanction, or to detect if the adulteration could cause health issues or not.

10.- BIBLIOGRAPHY

[1] J. Spink, W. Chen, G. Zhang, and C. Speier-Pero, "Introducing the Food Fraud Prevention Cycle (FFPC): A dynamic information management and strategic roadmap," *Food Control*, vol. 105, pp. 233–241, Nov. 2019, doi: 10.1016/J.FOODCONT.2019.06.002.

[2] E. K. Essuman, E. Teye, R. G. Dadzie, and L. K. Sam-Amoah, "Consumers' Knowledge of Food Adulteration and Commonly Used Methods of Detection," *J Food Qual*, vol. 2022, 2022, doi: 10.1155/2022/2421050.

[3] J. C. Moore, J. Spink, and M. Lipp, "Development and Application of a Database of Food Ingredient Fraud and Economically Motivated Adulteration from 1980 to 2010," *J Food Sci*, vol. 77, no. 4, pp. R118–R126, Apr. 2012, doi: 10.1111/J.1750-3841.2012.02657.X.

[4] F. Orlandi, F. Aguilera, C. Galán, M. Msallem, and M. Fornaciari, "OLIVE YIELDS FORECASTS AND OIL PRICE TRENDS IN MEDITERRANEAN AREAS: A COMPREHENSIVE ANALYSIS OF THE LAST TWO DECADES," *Exp Agric*, vol. 53, no. 1, pp. 71–83, Jan. 2017, doi: 10.1017/S0014479716000077.

[5] E. Casadei *et al.*, "Emerging trends in olive oil fraud and possible countermeasures," *Food Control*, vol. 124, p. 107902, Jun. 2021, doi: 10.1016/J.FOODCONT.2021.107902.

[6] R. E. Bartt and J. L. Topel, "Autoimmune and Inflammatory Disorders," *Textbook of Clinical Neurology: Third Edition*, pp. 1155–1184, 2007, doi: 10.1016/B978-141603618-0.10050-5.

[7] E. Gelpí *et al.*, "The Spanish toxic oil syndrome 20 years after its onset: a multidisciplinary review of scientific knowledge.," *Environ Health Perspect*, vol. 110, no. 5, p. 457, May 2002, doi: 10.1289/EHP.110-1240833.

[8] *Reg. (EU) 1308/2013*. UE, 2022. doi: 10.12.2023.

[9] L. Conte *et al.*, "Olive oil quality and authenticity: A review of current EU legislation, standards, relevant methods of analyses, their drawbacks and

recommendations for the future," *Trends Food Sci Technol*, vol. 105, pp. 483–493, Nov. 2020, doi: 10.1016/J.TIFS.2019.02.025.

[10] I. Durán Merás, J. Domínguez Manzano, D. Airado Rodríguez, and A. Muñoz de la Peña, "Detection and quantification of extra virgin olive oil adulteration by means of autofluorescence excitation-emission profiles combined with multi-way classification," *Talanta*, vol. 178, pp. 751–762, Feb. 2018, doi: 10.1016/J.TALANTA.2017.09.095.

[11] H. S. Green *et al.*, "A rapid method for the detection of extra virgin olive oil adulteration using UHPLC-CAD profiling of triacylglycerols and PCA," *Food Control*, vol. 107, Jan. 2020, doi: 10.1016/J.FOODCONT.2019.106773.

[12] T. K. de Lima, M. Musso, and D. Bertoldo Menezes, "Using Raman spectroscopy and an exponential equation approach to detect adulteration of olive oil with rapeseed and corn oil," *Food Chem*, vol. 333, Dec. 2020, doi: 10.1016/J.FOODCHEM.2020.127454.

[13] F. T. Borghi *et al.*, "Quantification and classification of vegetable oils in extra virgin olive oil samples using a portable near-infrared spectrometer associated with chemometrics," *Microchemical Journal*, vol. 159, Dec. 2020, doi: 10.1016/J.MICROC.2020.105544.

[14] A. Garrido-Varo, M. T. Sánchez, M. J. De la Haba, I. Torres, and D. Pérez-Marín, "Fast, Low-Cost and Non-Destructive Physico-Chemical Analysis of Virgin Olive Oils Using Near-Infrared Reflectance Spectroscopy," *Sensors 2017, Vol. 17, Page 2642*, vol. 17, no. 11, p. 2642, Nov. 2017, doi: 10.3390/S17112642.

[15] C. Melendreras *et al.*, "An Affordable NIR Spectroscopic System for Fraud Detection in Olive Oil," *Sensors*, vol. 23, no. 3, p. 1728, Feb. 2023, doi: 10.3390/s23031728.

[16] "COI/T.20/Doc. No 34- METHOD DETERMINATION OF FREE FATTY ACIDS, COLD METHOD." Accessed: May 21, 2024. [Online]. Available: http://www.internationaloliveoil.org/

[17] D. Alejandro, A. Forigua, and J. Meija, "Titration endpoint challenge", doi: 10.1007/s00216-018-1430-y.

[18] "COI/T.20/Doc. No 28- METHOD OF ANALYSIS DETERMINATION OF THE CONTENT OF WAXES, FATTY ACID METHYL ESTERS AND FATTY ACID ETHYL ESTERS BY CAPILLARY GAS CHROMATOGRAPHY." Accessed: May 21, 2024. [Online]. Available: http://www.internationaloliveoil.org/

[19] D. Harvey, *Modern analytical chemistry*. McGraw-Hill, 2000.

[20] T. Hayakawa *et al.*, "A Simple Screening Method for Extra Virgin Olive Oil Adulteration by Determining Squalene and Tyrosol," *J. Oleo Sci*, vol. 69, no. 7, pp. 677–684, 2020, doi: 10.5650/jos.ess20033.

[21] W. Minkina, "How Infrared Radiation Was Discovered—Range of This Discovery and Detailed, Unknown Information," *Applied Sciences 2021, Vol. 11, Page 9824*, vol. 11, no. 21, p. 9824, Oct. 2021, doi: 10.3390/APP11219824.

[22] Ciurczak Emil W. and Burns Donald A., "Handbook of near-infrared analysis / edited by Donald A. Burns, Emil W. Ciurczak," pp. xiv, 681 p. :, 1992, Accessed: May 18, 2024. [Online]. Available: https://books.google.com/books/about/Handbook_of_Near_infrared_Analysis.html?hl=es&id=fRfwAAAAMAAJ

[23] D. L. Pavia, G. M. Lampman, and G. S. Kriz, "Introduction to Spectroscopy BT - Introduction to Spectroscopy," *Introduction to Spectroscopy*, pp. 102–166, 2001, Accessed: May 19, 2024. [Online]. Available: papers3://publication/uuid/DA5E8589-68A0-4D2F-BB5E-7C12F3FA0CC3

[24] Y. Ozaki, C. Huck, S. Tsuchikawa, and S. B. Engelsen, *Near-Infrared Spectroscopy: Theory, Spectral Analysis, Instrumentation, and Applications*. Springer Singapore, 2020. doi: 10.1007/978-981-15-8648-4/COVER.

[25] N. Kumar, A. Bansal, G. S. Sarma, and R. K. Rawal, "Chemometrics tools used in analytical chemistry: An overview," *Talanta*, vol. 123, pp. 186–199, Jun. 2014, doi: 10.1016/J.TALANTA.2014.02.003.

[26] P. C. Jurs, B. R. Kowalski, T. L. Isenhour, and C. N. Reilley, "An Investigation of Combined Patterns from Diverse Analytical Data Using Computerized Learning Machines," *Anal Chem*, vol. 41, no. 14, pp. 1949–

1953, 1969, doi: 10.1021/AC50159A027/ASSET/AC50159A027.FP.PNG_V03.

[27] P. H. Weiner, E. R. Malinowski, and A. R. Levinstone, "Factor analysis of solvent shifts in proton magnetic resonance," *Journal of Physical Chemistry*, vol. 74, no. 26, pp. 4537–4542, 1970, doi: 10.1021/J100720A013/ASSET/J100720A013.FP.PNG_V03.

[28] D. L. Massart, C. Janssens, R. Smits, and L. Kaufman, "Application OT the i neory OT Graphs to the Optimization of unromatographic Separation Schemes for Multicomponent Samples," *Anal Chem*, vol. 44, no. 14, pp. 2390–2393, Dec. 1972, doi: 10.1021/AC60322A022/ASSET/AC60322A022.FP.PNG_V03.

[29] S. Wold, "Spline Functions in Data Analysis," *Technometrics*, vol. 16, no. 1, p. 1, Feb. 1974, doi: 10.2307/1267485.

[30] R. G. Brereton, "A short history of chemometrics: a personal view," *J Chemom*, vol. 28, no. 10, pp. 749–760, Oct. 2014, doi: 10.1002/CEM.2633.

[31] J. García Olmo, "Clasificación y autentificación de canales de cerdo ibérico mediante espectroscopía de infrarrojo cercano (NIRS)," 2009.

[32] M. Otto, *Chemometrics : statistics and computer application in analytical chemistry*. Wiley-VCH, 2017.

[33] H. Martens, "Multivariate Analysis of Quality. An Introduction," *Meas Sci Technol*, Jan. 2001, Accessed: May 19, 2024. [Online]. Available: https://www.academia.edu/60917453/Multivariate_Analysis_of_Quality_An_Introduction

[34] P. Oliveri, C. Malegori, and M. Casale, "Chemometrics: multivariate analysis of chemical data," *Chemical Analysis of Food: Techniques and Applications, Second Edition*, pp. 33–76, Jan. 2020, doi: 10.1016/B978-0-12-813266-1.00002-4.

[35] S. Wold and M. Sjöström, "Chemometrics, present and future success," *Chemometrics and Intelligent Laboratory Systems*, vol. 44, no. 1–2, pp. 3–14, Dec. 1998, doi: 10.1016/S0169-7439(98)00075-6.

[36] L. C. Lee, C. Y. Liong, K. Osman, and A. A. Jemain, "Effects of scatter-correction pre-processing methods and spectral derivative algorithms on forensic classification of paper," *AIP Conf Proc*, vol. 1750, no. 1, p. 060013, Jun. 2016, doi: 10.1063/1.4954618.

[37] J. Kobina, M. Biney, J. Ruth Blöcher, L. Borůvka, and R. Vašát, "Does the limited use of orthogonal signal correction pre-treatment approach to improve the prediction accuracy of soil organic carbon need attention?," 2021, doi: 10.1016/j.geoderma.2021.114945.

[38] F. Westad, "A retrospective look at cross model validation and its applicability in vibrational spectroscopy," 2021, doi: 10.1016/j.saa.2021.119676.

[39] R. J. Barnes, M. S. Dhanoa, and S. J. Lister, "Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra," *Appl Spectrosc*, vol. 43, no. 5, pp. 772–777, 1989, doi: 10.1366/0003702894202201.

[40] H. Jonsson and J. Gabrielsson, "Evaluation of Preprocessing Methods," *Comprehensive Chemometrics*, vol. 2, pp. 199–206, 2009, doi: 10.1016/B978-044452701-1.00112-5.

[41] M. S. Dhanoa, R. Sanderson, S. J. Lister, and R. J. Barnes, "The Link between Multiplicative Scatter Correction (MSC) and Standard Normal Variate (SNV) Transformations of NIR Spectra," *Journal of Near Infrared Spectroscopy, Vol. 2, Issue 1, pp. 43-47*, vol. 2, no. 1, pp. 43–47, Jan. 1994, Accessed: May 20, 2024. [Online]. Available: https://opg.optica.org/abstract.cfm?uri=jnirs-2-1-43

[42] L. E. Agelet and C. R. Hurburgh, "A Tutorial on Near Infrared Spectroscopy and Its Calibration," *Crit Rev Anal Chem*, vol. 40, no. 4, pp. 246–260, Oct. 2010, doi: 10.1080/10408347.2010.515468.

[43] N. B. Gallagher, "Savitzky-Golay Smoothing and Differentiation Filter", Accessed: May 20, 2024. [Online]. Available: www.Eigenvector.com

[44] A. Savitzky and M. J. E. Golay, "Smoothing and Differentiation of Data by Simplified Least Squares Procedures," *Anal Chem*, vol. 36, no. 8, pp. 1627–

1639, Jul. 1964, doi: 10.1021/AC60214A047/ASSET/AC60214A047.FP.PNG_V03.

[45] W. H. Press, "Numerical recipes in C : the art of scientific computing," p. 994, 1992, Accessed: May 20, 2024. [Online]. Available: https://search.worldcat.org/title/1010839606

[46] I. Chandrasekaran, S. S. Panigrahi, L. Ravikanth, and C. B. Singh, "Potential of Near-Infrared (NIR) Spectroscopy and Hyperspectral Imaging for Quality and Safety Assessment of Fruits: an Overview," *Food Anal Methods*, vol. 12, no. 11, pp. 2438–2458, Nov. 2019, doi: 10.1007/S12161-019-01609-1.

[47] S. Sirek and R. Kane, "The Tungsten Halogen Lamp," *Revolution in Lamps*, pp. 75–91, Dec. 2020, doi: 10.1201/9781003150985-4.

[48] A. A. Ismail, F. R. van de Voort, and J. Sedman, "Chapter 4 Fourier transform infrared spectroscopy: Principles and applications," *Techniques and Instrumentation in Analytical Chemistry*, vol. 18, no. C, pp. 93–139, Jan. 1997, doi: 10.1016/S0167-9244(97)80013-3.

[49] J. Workman, M. Koch, B. Lavine, and R. Chrisman, "Process analytical chemistry," *Anal Chem*, vol. 81, no. 12, pp. 4623–4643, Jun. 2009, doi: 10.1021/AC900778Y/ASSET/AC900778Y.FP.PNG_V03.

[50] I. Zafra Chillerón, "Potencial de datos espectrales NIRS 'on-site' para la detección de desviaciones de producto en leche respecto al estándar de calidad y seguridad," 2014. Accessed: May 31, 2024. [Online]. Available: https://digibuo.uniovi.es/dspace/handle/10651/27953

[51] W. ; Zhang *et al.*, "A Review of Machine Learning for Near-Infrared Spectroscopy," *Sensors 2022, Vol. 22, Page 9764*, vol. 22, no. 24, p. 9764, Dec. 2022, doi: 10.3390/S22249764.

[52] A. Saleem, "NIR spectroscopy for personal screening," 2011.

[53] Y. Pu, D. Pérez-Marín, N. O'shea, and A. Garrido-Varo, "Recent Advances in Portable and Handheld NIR Spectrometers and Applications in Milk, Cheese and Dairy Powders," *Foods 2021, Vol. 10, Page 2377*, vol. 10, no. 10, p. 2377, Oct. 2021, doi: 10.3390/FOODS10102377.

[54] C. Tang *et al.*, "A review of graphene-based temperature sensors," *Microelectron Eng*, vol. 278, Jun. 2023, doi: 10.1016/j.mee.2023.112015.

[55] R. Sczupak *et al.*, "Measurement and alignment of linear variable filters," *Photonic Instrumentation Engineering V*, p. 105390, 2018, doi: 10.1117/12.2290689.

[56] M. Hernández-Jiménez, M. I. González-Martín, I. Martínez-Martín, I. Revilla, and A. M. Vivar-Quintana, "Carbon stable isotopes, fatty acids and the use of NIRS to differentiate IBERIAN pigs," *Meat Sci*, vol. 182, p. 108619, Dec. 2021, doi: 10.1016/J.MEATSCI.2021.108619.

[57] J. A. Abecia *et al.*, "Sensores NIRS portátiles para el control de calidad individualizado de jamones de cerdo ibérico antes y después de la salazón," *XVI Jornadas sobre Producción Animal: 19 y 20 de mayo de 2015, Zaragoza, Vol. 2, 2015, ISBN 978-84-606-7969-1, págs. 603-605*, vol. Tomo I, pp. 603–605, 2015, Accessed: May 21, 2024. [Online]. Available: https://dialnet.unirioja.es/servlet/articulo?codigo=8658146&info=resumen&idioma=ENG

[58] D. Tejerina, S. García Torres, M. Cabeza de Vaca, A. Ortiz, and M. P. Romero, "Evaluation of near-infrared spectroscopy (NIRS) for the quality control of packaged cured ham-sliced from Iberian pigs," *Archivos de zootecnia, ISSN-e 1885-4494, ISSN 0004-0592, Nº Extra 1, 2018, págs. 231-234*, no. 1, pp. 231–234, 2018, Accessed: May 21, 2024. [Online]. Available: https://dialnet.unirioja.es/servlet/articulo?codigo=6537156&info=resumen&idioma=SPA

[59] *Listado de denominaciones de origen protegidas (DOP) e indicaciones geográficas protegidas (IGP) españolas registradas en la Unión Europea*. 2023.

[60] M. J. De la Haba, M. Arias, P. Ramírez, M. I. López, and M. T. Sánchez, "Characterizing and Authenticating Montilla-Moriles PDO Vinegars Using Near Infrared Reflectance Spectroscopy (NIRS) Technology," *Sensors (Basel)*, vol. 14, no. 2, p. 3528, Feb. 2014, doi: 10.3390/S140203528.

[61] M. del R. Márquez Ortega, "Aseguramiento de la calidad y autentificación de vinagres de vino mediante sensores nirs," p. 1, 2020, Accessed: May 21, 2024. [Online]. Available: https://dialnet.unirioja.es/servlet/tesis?codigo=294018&info=resumen&idioma=SPA

[62] R. Ríos-Reina, D. L. García-González, R. M. Callejón, and J. M. Amigo, "NIR spectroscopy and chemometrics for the typification of Spanish wine vinegars with a protected designation of origin," *Food Control*, vol. 89, pp. 108–116, Jul. 2018, doi: 10.1016/J.FOODCONT.2018.01.031.

[63] V. Giannetti, M. B. Mariani, and P. Mannino, "Characterization of the Authenticity of Pasta di Gragnano Protected Geographical Indication Through Flavor Component Analysis by Gas Chromatography-Mass Spectrometry and Chemometric Tools," *J AOAC Int*, vol. 99, no. 5, p. 1279, 2016, doi: 10.5740/jaoacint.16-0137.

[64] P. Firmani, G. La Piscopia, R. Bucci, F. Marini, and A. Biancolillo, "Authentication of P.G.I. Gragnano pasta by near infrared (NIR) spectroscopy and chemometrics," *Microchemical Journal*, vol. 152, p. 104339, Jan. 2020, doi: 10.1016/J.MICROC.2019.104339.

[65] A. Biancolillo *et al.*, "Authentication of an Italian PDO hazelnut ('Nocciola Romana') by NIR spectroscopy," *Environ Sci Pollut Res Int*, vol. 25, no. 29, pp. 28780–28786, Oct. 2018, doi: 10.1007/S11356-018-1755-2.

[66] M. Tarapoulouzi *et al.*, "A Pre-Trial Study to Identify Species of Origin in Halloumi Cheese Utilising Chemometrics with Near-Infrared and Hyperspectral Imaging Technologies," *Analytica 2024, Vol. 5, Pages 17-27*, vol. 5, no. 1, pp. 17–27, Jan. 2024, doi: 10.3390/ANALYTICA5010002.

[67] L. Vasconcelos *et al.*, "Contribution to Characterizing the Meat Quality of Protected Designation of Origin Serrana and Preta de Montesinho Kids Using the Near-Infrared Reflectance Methodology," *Foods 2024, Vol. 13, Page 1581*, vol. 13, no. 10, p. 1581, May 2024, doi: 10.3390/FOODS13101581.

[68] J. Arruabarrena Gamboa, "Espectroscopía NIR y Raman para el control de calidad en industria farmacéutica," 2014.

[69] "Flow through cuvette (1767008510-40)." Accessed: Jun. 18, 2024. [Online]. Available: https://www.hellma.com/en/laboratory-supplies/cuvettes/cuvette-finder/cuvette/1767008510-40

[70] "DLP ® NIRscan™ Nano EVM User's Guide User's Guide," 2015. [Online]. Available: www.ti.com

[71] E. Pruett, "Application Report Texas Instruments DLP ® Spectrometer Design Considerations," 2014, Accessed: May 30, 2024. [Online]. Available: www.ti.com

[72] L. G. Weyer and S.-C. Lo, "Spectra-Structure Correlations in the Near-infrared," 2006, doi: 10.1002/9780470027325.s4102.

[73] A. Arroyo-Cerezo, X. Yang, A. M. Jiménez-Carvelo, M. Pellegrino, A. Felicita Savino, and P. Berzaghi, "Assessment of extra virgin olive oil quality by miniaturized near infrared instruments in a rapid and non-destructive procedure," *Food Chem*, vol. 430, p. 137043, Jan. 2024, doi: 10.1016/J.FOODCHEM.2023.137043.

[74] N. Camino Pérez, "Caracterización del perfil de ácidos grasos en aceites vegetales mediante HPLC," 2021.

ANNEX I

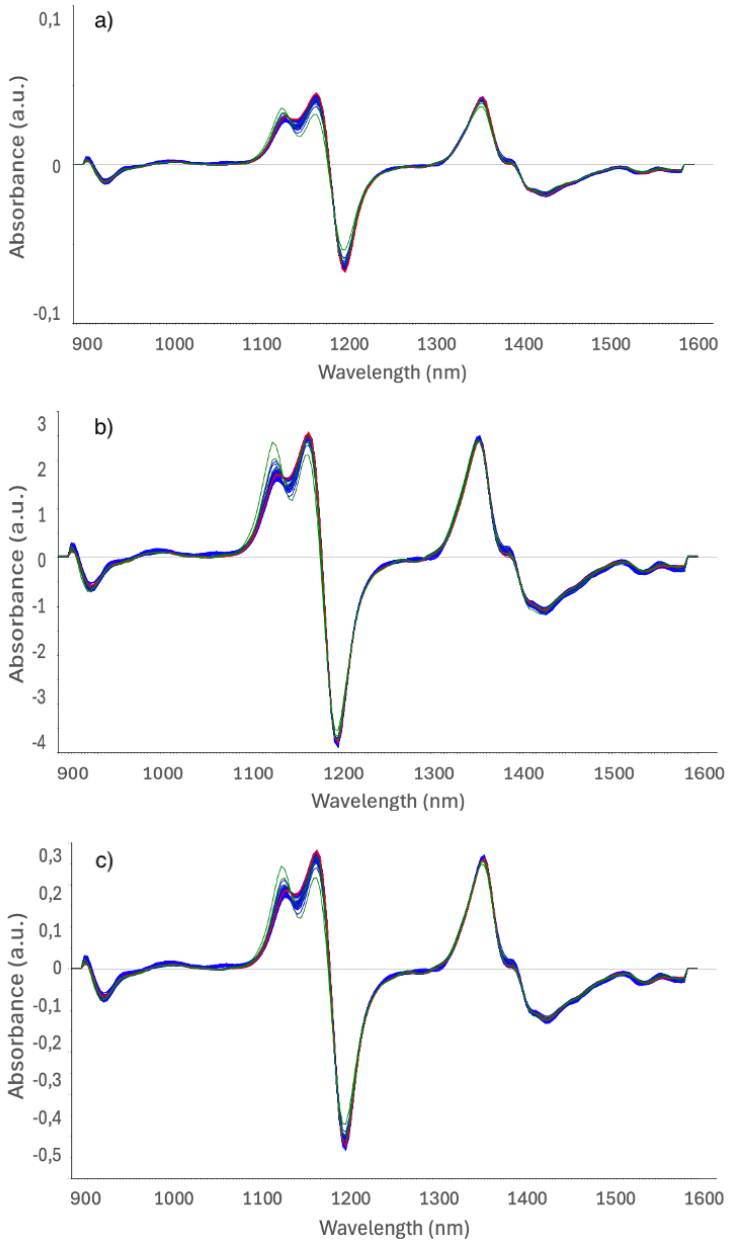

Figure 11.1- Spectral plots after chemometric treatments with first order derivatives (with no seeds oil). a) SG-1-2-4-4. b) SG-1-2-4-4 + SNV. c) SNV + SG-1-2-4-4.

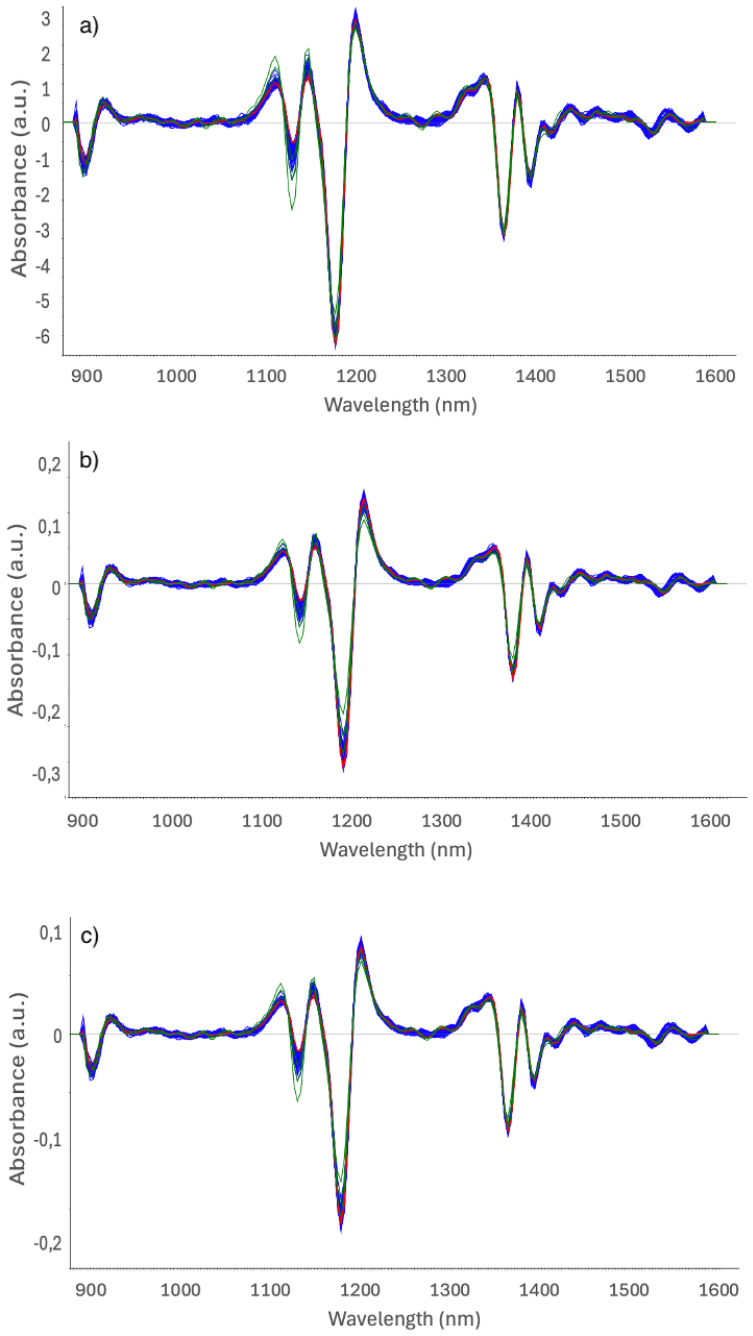

Figure 11.2- Spectral plots after chemometric treatments with second order derivatives (with no seeds oil). a) SG-2-2-4-4. b) SG-2-2-4-4 + SNV. c) SNV + SG-2-2-4-4.

ANNEX II

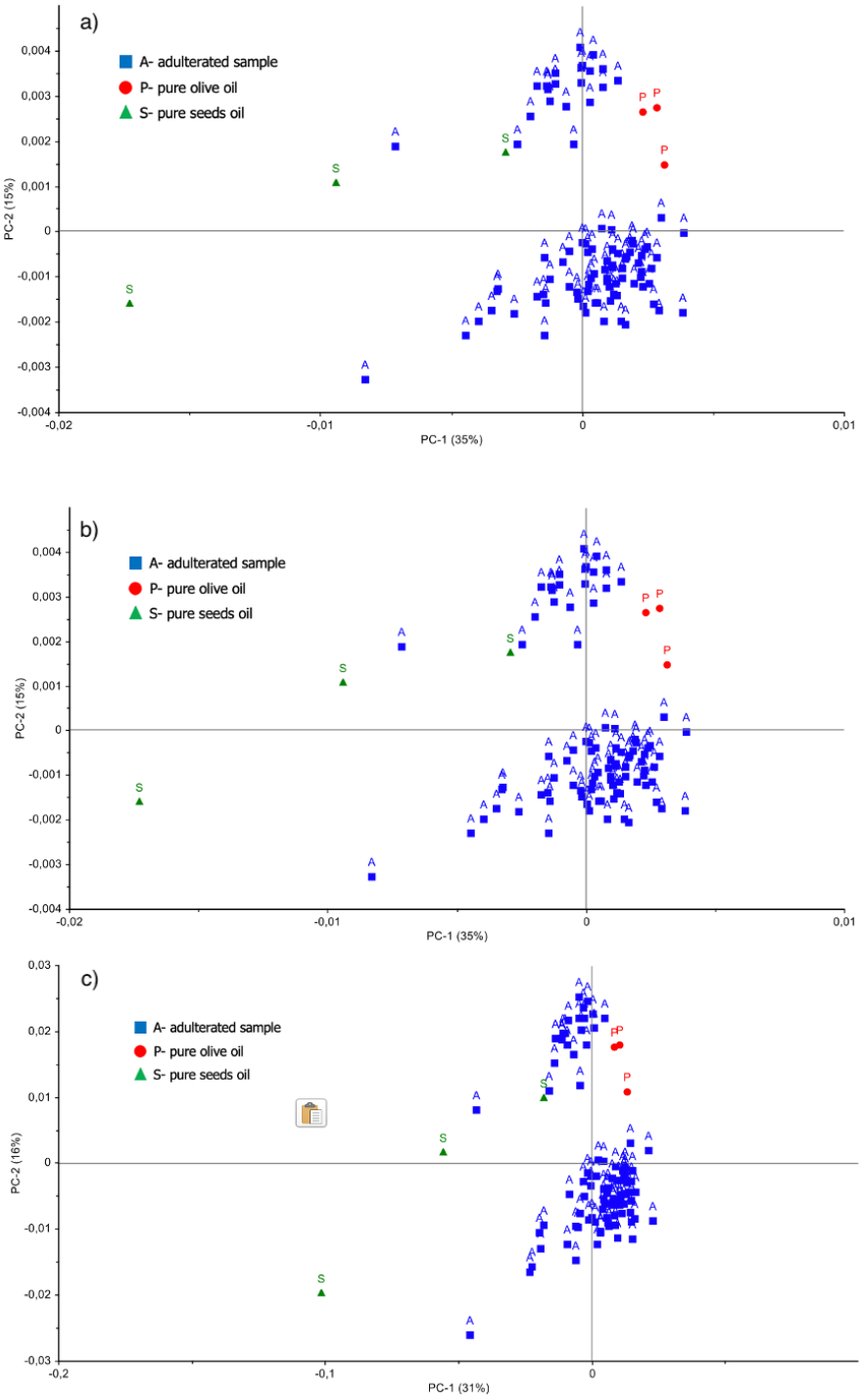

Figure 12.1- PCA plots of chemometric treatments with second order derivatives (with no seeds oil). a) SG-2-2-4-4. b) SG-2-2-4-4 + SNV. c) SNV + SG-2-2-4-4.