Verdú-Amat, S.; Fuentes López, C.; Fuentes López, A.; Pérez, AJ.; Barat Baviera, JM.; Grau Meló, R. (2024). Control of cod-liver oil composition with laser scattering imaging combined with machine learning procedures: The cases of adulteration and oxidation. Journal of Food Engineering. 370. https://doi.org/10.1016/j.jfoodeng.2024.111955

Control of cod-liver oil composition with RGB-laser scattering imaging combined with machine learning procedures: the cases of adulteration and oxidation

Samuel Verdú[a], Cristina Fuentes[a], Ana Fuentes[a], Alberto J. Pérez[b], José M. Barat[a], Raúl Grau[a]

[a]Dpto. de Tecnología de alimentos, Universitat Politècnica de Valencia

[b]Dpto. de Informática de Sistemas y Computadores, Universitat Politècnica de Valencia

*Author for correspondence: Samuel Verdú

Address: Edificio 8E - Acceso F – Planta 0

Ciudad Politécnica de la Innovación

Universidad Politécnica de Valencia

Camino de Vera, s/n

46022 VALENCIA – SPAIN

E-mail: saveram@upvnet.upv.es

Abstract

Chemical changes in cod-liver oil produced by oxidation and adulterations with other oils were modelled using RGB-laser scattering imaging. Two types of composition-altered cod-liver oil were: oxidised oil at three different temperatures (4, 20, and 40ºC) and cod-liver oil adultered with wheat-germ, soybeana, sesame and corn. Both types of altered oils and control samples were analyzed by the imaging technique and chemical measurements were also carried out to know the oxidation status. The capacity of the technique for detecting pure cod-liver oil was tested by applying multivariate regression and discriminant procedures based on PLS using the information captured from each type of laser (650nm, 550nm and 450nm). The results showed the capacity of the technique to capture the variability provided by the cod-liver oil against the other vegetable oils used as adulterants. It was discriminate from all vegetable oils and all them were detected when were added as adulterants. Changes in the oxidation status were also modelled by predicting the oxidation parameters with $R^2>0.90$, independently of the temperature of storage. Those capacities made it possible to discriminate pure cod-liver from all tested adulterated-oxidised samples performing a common model. The prediction capacity was synergic when models were performed including the three lasers, as opposed to the results obtained with singly-laser models.

Keywords: laser scattering imaging, cod-liver oil, oxidation, adulteration, non-destructive inspection

1. Introduction

The beneficial effects of replacing dietary saturated fatty (SFAs) acids with unsaturated fatty acids (UFAs) on human health have been observed through extensive research at all levels of the hierarchy of evidence: in vitro, in vivo, cohort studies, meta-analysis, etc. Most of them are consistent in that reducing dietary SFAs while UFAS increases LDL cholesterol and cardiovascular disease risk significantly reduced (Sacks et al., 2017). There are diverse sources of UFAs in the human diet. Those compounds are present naturally in raw animal, vegetal and algae-origin materials. Their proportion and composition vary depending on the properties of the original organism and how they are processed to derive into a food product. The sources of UFAs could form part of a human diet as part of the original raw material matrices (fresh mackerel, almonds, avocado, etc.), as extracted oils (cod-liver oil, sesame oil, soybean oil, etc.) or as part of a processed product formula (pre-cooked meal, cereal derivate products, dairy foods, etc.). Thus, these types of fats represent added-value products from which enrich the diet as supplements or ingredients.

One of the most valuable sources is the liver of the cod. The cod-liver oil (CO) represents a traditional source of healthy compounds such as vitamins A, D and omega-3 fatty acids of (eicosapentaenoic (EPA) and docosahexaenoic (DHA) acids) (Cortese et al., 2015). The results reported from epidemiological studies revealed that high fish oil consumption is directly related to low mortality following coronary heart disease and some types of cancer (Lentjes et al., 2014). Those factors stimulate the demand of this oil and then increase the price value in the market. Therefore, CO fraud with low price oils such deteriorated batches and vegetables oils can be done in order to increase economical benefits. From a technical and quality control points of view, the adulterations in oils are more difficult to detect when the composition of the adulterant is chemically near to the composition of the authentic one.

Some techniques to detect different adulterations and oxidation status in CO has been reported previously, which are based in different technologies. Part of the power of those techniques lead in their combination with machine learning and multivariate statistical methods to process the

obtained data. The detected oils had both vegetal and animal origin. Some examples in the detection of changes produced by vegetable oils are canola, sunflowers, corn, soybean, and walnut oils using FTIR combined, nuclear magnetic resonance spectroscopy, gas chromatography and Raman spectroscopy combined with partial least squares (PLS), principal component regression (PCR) , discriminant analysis (DA), artificial neural networks (ANN), etc. (Ahmmed et al., 2022; Giese et al., 2019; Abdul Rohman & Che Man, 2011a). In the case of adulterant oil from animal origin, chicken, beef and mutton fat were successfully detected following the same procedures (A. Rohman & Che Man, 2011; Abdul Rohman & Che Man, 2011b). In the other hand, some of the most used techniques to study the deterioration of CO in oxidation terms are the characterisation of the peroxide value and anisidine by the AOCS Official Methods Cd 8b-90 and Cd 18-90, respectively (AOCS, 2017).

One family of spectroscopic techniques that works fast, collecting a significant amount of chemical information on which to apply machine learning procedures, is imaging analysis. These methods capture spectral information from interactions between light and matter in a two-dimensioned sensor, which is divided into pixels. The pixels are the unit of signal detection, from which a map of the surface of a product could be reconstructed using the spectral information. The use of this type of technique in characterisation of CO is limited, however, its suitability to study the quality of similar edible oils has been reported with techniques such hyperspectral imaging. Some examples are the categorisation of sesame oils from other adulterated oils (Xie et al., 2014), evaluation of various quality attributes such as peroxide index, acidity of olive oils and prevent its adulteration from other low-grade oils (Cano Marchal et al., 2013). In this sense, the laser-scattering imaging technique has not been tested in that type of food matrix, however it has great potential as tool to characterizing changes in oils, concretely in CO, since it can be combined with machine learning procedures to analysing the generated data. This technique is based on capturing digital images of the diffraction patterns generated by the interaction between a given food matrix and a laser light. Thus, the coherent light is transmitted through the material up to the

surface, where it has given scattered properties depending on the sample's structure and components. These phenomena depend on the physicochemical properties of the matrix, as backscattered photons inherently interact with internal components (Mollazade et al., 2013). Therefore, the suitability to be applied on CO inspection is founded in the results obtained by this technique modelling physicochemical properties of different food matrix such as the moisture content of peppers during drying (Romano et al., 2012), the textural properties of vegetable creams (Verdú et al., 2018), the effect of fibre enrichment on biscuit properties (Verdú et al., 2019), or characterising the curding process of milk with different fat content (Verdú et al., 2021).

Then, in this work, the laser-backscattering imaging technique is proposed as a control technique of cod-liver oil in terms of adulteration and oxidative status detection. The study focused on the non-destructive prediction of the type of adulterants vegetable oils and oxidation parameters testing the imaging information from three different wavelength lasers and the combination of all of them applying machine learning procedures.

2. Material and Methods

2.1 Experimental procedure

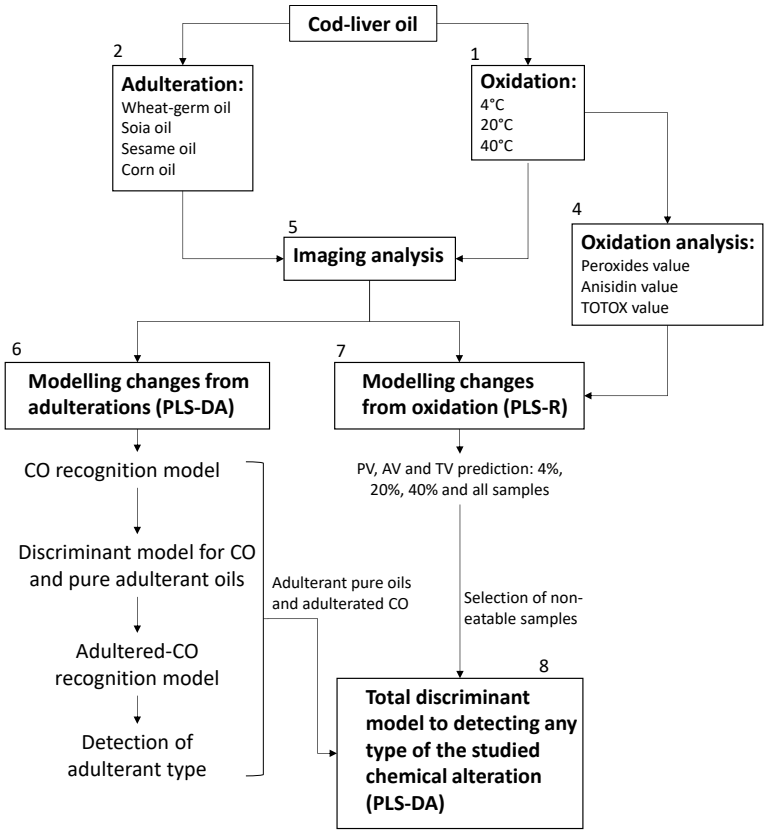

Figure 1. Scheme of the experimental phases. CO: cod-liver oil.

Modelling of the chemical changes in cod-liver oil (CO) due to oxidation processed and adulterations with other oils using RGB-laser scattering imaging was carried out following the phases showed in the Figure 1. Two types of composition-altered cod-liver oil were prepared. The first type was prepared based on allowing oxidation at three different temperatures (40, 20, and 4 ^{o}C), for 17, 63 and 105 days, respectively (Figure 1-1). The second type was generated by adulterating cod-liver oil with vegetable oils (wheat-germ, soybean, sesame and corn) (Figure 1-2). Both types of altered oils and control samples were analyzed by the imaging technique (Figure 1-5). In the case of the oxidized oils, chemical measurements were also carried out to know the oxidation status of the samples (peroxide value, anisidine value and total oxidation value, Figure 1-4)). The capacity of the technique for discriminating CO from other oils and detecting adulterations was tested by applying multivariate discriminant procedures (Figure 1-6). In the same way, the capacity to predict the CO oxidation status, multivariate regression procedures were applied to model the oxidation analysed parameters (Figure 1-7). Finally, a common

discriminant model was made aimed at detecting any type of the studied chemical alteration at once (Figure 1-8).

2.2 Chemicals and raw materials/ Reagents

All the reagents were of standard analytical grade. Iodine, potassium iodide, potassium iodate, sodium thiosulfate, and sulfuric acid were supplied by Scharlab (Spain). 1-decanol was acquired from Thermo Fisher Scientific (USA). Glacial acetic acid was purchased from Labkem (Spain) and *p*-anisidine was obtained by Sigma-Aldrich (Spain).

2.3 Peroxide value

Primary oxidation products were assessed by PV measurement using the redox-potentiometric titration method. Approximately 1 g of oil was dissolved in 10 mL of a glacial acetic acid and 1-decanol solution (3:2 v/v, containing 0.005 % iodine). Then, 200 μL of a saturated KI solution were added. After 1 min incubation in darkness, 50 mL of distilled water were added, and the mixture was titrated with a 0.01 M sodium thiosulfate solution using an automated titration instrument Metrohm 905 titrando (Herisau, Suiza). The PV was calculated according to the following equation:

$$PV\ (mEq/Kg) = \ 10 \times (Vs - Vb) \times (\frac{N}{W})$$

where Vs is the volume of thiosulfate consumed for each sample (mL), Vb is the volume of thiosulfate consumed for the blank (mL), N is the titer of the thiosulfate solution and W is the weight of the sample (g).

2.4 Anisidine value

Secondary oxidation products were assessed by AV measurement using the American Oil Chemists' Society (AOCS) Official Method Cd 18–90. Briefly, 1 mL of oil sample was mixed with 25 mL isooctane and centrifugated at 4000 rpm and 15 ºC for 20 min. Then, 2 mL of the supernatant were mixed with 400 μL of the p-anisidine solution (0.25 % *p*-anisidine in glacial acetic acid) and incubated at room temperature for 10 min in darkness. The absorbance of the supernatant was measured using a Spectrophotometer (Helios Zeta UV-VIS, Thermo Fisher

Scientific, USA) at a wavelength of 350 nm before (A1 value) and after incubation (A2 value). The AV value for each sample was calculated according to the following formula:

$$AV = 25\ (1.2A2 - A1)$$

2.5 Totox value

The total oxidation of the samples was calculated as the total oxidation value (Totox) according to the formula:

$$TV = 2PV + AV$$

2.6 Oils

The cod-liver oil (CO) and the vegetable oils used as adulterants (corn (C), sesame (Se), wheatgerm (WG) and soybean (So)) were obtained from Guinama S.L.U, Spain. Table 1 shows physicochemical characterisation facilitated by the provider. The adulterated versions of CO were made by adding 1 and 5 % vegetable oils. Twenty samples of CO per type of adulterant oil were produced.

Table 1. Physicochemical information of cod-liver oil and adulterant oils.

Property	***CO***		***Se***		***So***		***WG***		***C***	
Density (g/mL)	0.923	(0.006)	0.9	(0.007)	0.917	(0.005)	0.9	(0.005)	0.9	(0.005)
Acid value (mg KOH/g)	0.15	(0.71)	0.2	(0.35)	0.04	(0.01)	0.1	(0.01)	0.2	(0.5)
Refractive index	1.477	(0.005)	1.5	(0.005)	1.472	(0.004)	1.5	(0.005)	1.5	(0.005)
Palmitic acid (C16:0) (%)	9.7	(1.9)	13	(3.2)	10.7	(1.9)	13	(1.1)	0.5	(0.2)
Estearic acid (C18:0) (%)	2.2	(0.1)	5.8	(1.3)	4	(1.1)	2.5	(0.8)	1.7	(1)
Oleic acid (C18:1) (%)	15.5	(2.4)	42	(3.8)	21.9	(3.2)	25	(2.8)	32	(2.9)
Linoleic acid (C18:2) (%)	1.4	(0.8)	38	(4.3)	54.4	(4.1)	6.4	(1.7)	53	(8.6)
Linolenic acid (C18:3) (%)	0.8	(0.2)	0.5	(0.2)	8.1	(1.3)	52	(5.3)	1.2	(0.5)

cod-liver oil (CO), corn oil(C), sesame oil (Se), wheatgerm oil (WG) and soybean oil (So). Values are expressed as mean (SD).

For studying the oxidation status of the oils, aliquots (20 mL) of fish oil were stored in darkness in 20-mL amber glass vials (28 mm diameter and 60 mm heigh) at 40, 20, and 4 ºC, for 17, 63 and 105 days, respectively. For each temperature, three vials were taken from the incubator at scheduled times, then PV and AV were analysed and the Totox value was calculated.

2.7 Imaging device

The imaging device was based on (Verdú et al., 2020). It was designed to capture, in digital images, the variability projected in the diffraction patterns generated after a light-food matrix interaction. The structured light (laser) transmits throughout the studied matrix, from bottom to upper surfaces, responding to possible variability sources due to interactions with the oil components. That variability could be modelled by data extracted from digital images of the resulting diffraction patterns. Figure 2 shows a scheme of the device setup. The elements were a digital camera, three laser lights, and a computer. The device was installed inside a dark box away from external light pollution. The camera's capture parameters were set to manual mode, avoiding uncontrolled alterations to the work regime (gain, shutter speed, white balance, etc.). A digital Logitech C920 camera (1080p/30 fps - 720p/30 fps) was the capture system. Images were acquired in the RGB format and were saved as JPEG (1980 × 1080). The camera was vertically placed 15 cm over the sample container located in the middle of the visual field. Used laser pointers (1mmθ, 50 mW) were three, which had the wavelengths 650 nm (R), 550nm (G) and 450nm (B). They were perpendicularly placed 20 cm under the zone where the wells plate was placed. Images of diffraction patterns from each wavelength were captured individually during light emission times, 3s per laser in the order R, G and B (Figure 1). The objective of this configuration was to transmit enough light to project the variability of the studied matrix without saturating the camera sensor. Seven replicates of images per sample were taken. Thus, 21 diffraction patterns per sample were captured in images to be later processed (7xR+7xG+7xB).

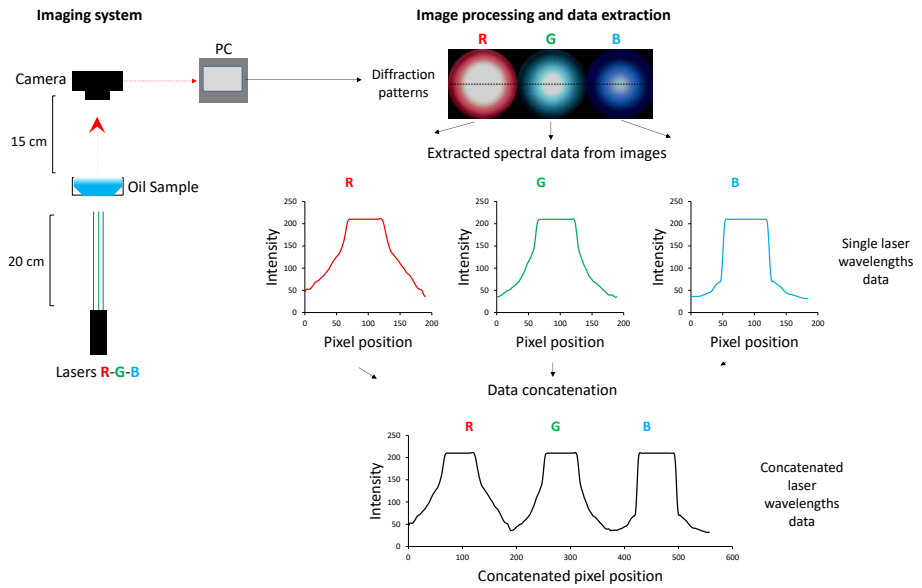

Figure 2. Scheme of the device and imaging data processing.

2.8 Image processing and data extraction

Images were processed after the capture procedure to extract information from the diffraction patterns generated during light-matrix interactions. The imaging processing was done with the aim of testing the information generated by each wavelength in an individual and concatenated mode. The image processing and data extraction was done as follows (Figure 2):

- Intensity from each pixel position placed across the diameter of each diffraction pattern was extracted in grayscale as a numeric vector formed by 195 values. This dataset was destined to test the prediction models by using the singly laser wavelengths.
- The extracted vectors from the R, G and B diffraction patterns were joined to create a concatenated version of the all lasers data (195 values x 3). This dataset was destined to create the prediction models by using the information from the three laser wavelengths at the same time.
- The final spectrum for each sample was the mean from the seven abovementioned measurements. That spectrum represented a fingerprint of 585 variables from each sample.

- The extracted spectra from the experiment were placed in matrices of *m* rows (number of samples) and *n*=195 columns (singly laser wavelength data studies) or n=585 columns (concatenated laser wavelength data studies). The format of the generated data matrix allowed for applying multivariate statistical methods to model the captured variance in the images.

2.9 Statistical analysis

The prediction models were carried out using Partial Least Squares procedures (PLS). That method projects variance to a reduced space of variance which explains the relationship between the two sets of variables X (imaging dataset) and Y (physicochemical properties), by maximising the covariance between datasets. It generates new synthetic variables, called latent variables (LVs), which collect the original variance from data in a reduced dimensional space. In this experiment, two type of PLS was used in function of the type of modelled response variable; numerical (oxidation status) or categorical (adulteration). Partial Least Square Regression (PLS-R) (Geladi & Kowalski, 1986), was used to generate linear regression models between oxidation parameters (PV, AV and TV) and imaging datasets to predict the oxidation status from the extracted imaging information. The regression models were evaluated based on the determination coefficient R^2 between measured and predicted datasets. Moreover, Partial Least Square Discriminant Analysis (PLS-DA) was used to generate classification models for differentiating fresh cod-liver from adulterated samples, non-eatable oxidized samples and other types of oils. The discriminant models were evaluated by F-score, calculated as follows:

$$F - score = 2x\frac{precision \cdot recall}{precision + recall} \quad (1)$$

Were precision and recall were calculated as follows:

$$precision = \frac{TP}{TP+FP} \quad (2)$$

$$recall = \frac{TP}{TP+FN} \quad (3)$$

TP, TN, FP and FN indicate true positives, true negatives, false positives and false negatives, respectively.

The performed discrimination models were:

- CO recognition model (CO-M): tested the capacity to discriminate CO from the remaining oils as a unique category (WG, Se, C and So). The stablished categories were two: CO and non-CO oils.
- Discriminant model for CO and pure adulterants (OT): aimed to test the capacity of the technique to capture the variance generated from each type of oil. The stablished categories were five: CO, WG, Se, C and So.
- Adulterated-CO recognition model (A-CO): aimed to discriminate pure CO from adulterated CO with any vegetable oil. The stablished categories were two: CO and adulterated CO.
- Detection of adulterant type model (OT-A-CO): aimed to detect adulterated CO and the type of adulterant added. The stablished categories were five: CO, WG, Se, C and So.
- Total discriminant model for CO (T-CO): aimed to discriminate CO from the rest of the studied oils, adulterated samples and non-eatable CO samples because its oxidation status. The limit to classify CO oil as non-eatable was 5 meq/kg O_2 for PV. Higher values are considered as non-acceptable for sensory attributes (Joint FAO/WHO Codex Alimentarius Commission, 2017). The stablished categories were two: pure CO and unsuitable oil (pure adulterant oils+ adulterated CO + oxidized CO).

The models were performed using 66 % samples as training lot by applying cross-validation procedures (venetian blinds) and validated using the remaining 33 % samples (test lot). That procedure was repeated with 100 different randomised training and test lots of samples. The optimum number of latent variables (NLV) were selected based on the minimum root means square error (RMSE) for each combination. Thus, the figures of merit for each prediction model

were expressed as the mean R^2 and F-score values from those 100 model replicates. Significance of the model differences was evaluated by ANOVA. Each type of model was performed from information on each single laser and the combination of them after apply a mean-centring process. Thus, four different models were tested for each application: R-models, G-models, B-models and R+G+B models (Figure 2).

3. Results and discussion

3.1 Analysis of oxidation status

Changes in the state of oxidation of fish oil during the storage at different temperatures are shown in Figure 3. As can be observed, the different deterioration parameters varied with storage temperature. At the beginning of the experiment the PV was 2.55 mEq O_2/kg. This value constantly increased for the 4 and 20 ºC samples during all the incubation period, while for 4 ºC samples a rapid increase, was followed by a stationary phase reached at eleven days of storage. CODEX/FAO has set limits for the quality parameters of fish oils intended for human consumption (Joint FAO/WHO Codex Alimentarius Commission, 2017). These regulatory upper limits are set at 5 mEq O_2/kg, 20, and 26, for PV, AV and Totox, respectively. After 2 days of incubation, the samples stored at 40 ºC exceeded the 5 mEq O_2/kg limit (5.91 mEq O_2/kg) stablished for PV. For 20 and 4 ºC, this value was exceeded after 14 days (6.14 mEq O_2/kg) and 31 days (7.99 mEq O_2/kg) of incubation, respectively. The AV of the samples increased with time for all the storage temperatures; however, only samples stored at 20 ºC (21.84) exceeded the 20 upper limit stablished for AV. Finally, Totox value of 4, 20, and 40 ºC samples, exceed the regulatory limit ($\leq$ 26) after 17, 63 and 105 days of incubation, respectively.

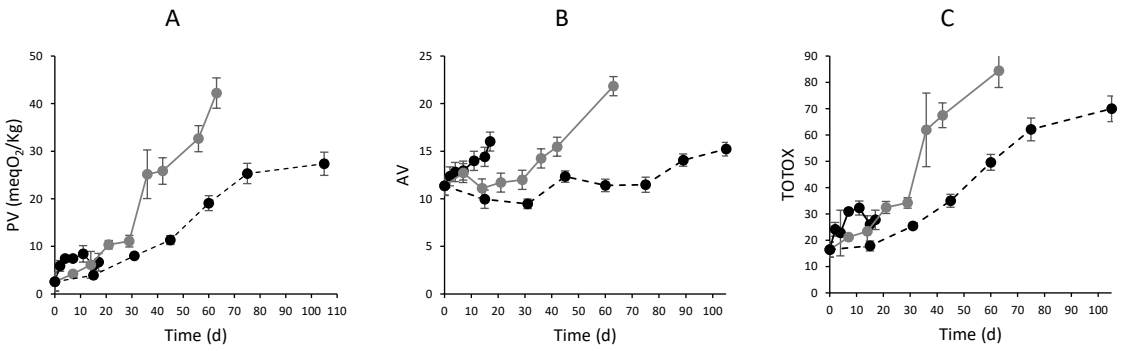

Figure 3. Evolution of the oxidation parameters peroxide value (A), anisidine value (B) and total oxidation value (C) for cod-liver oil at 4ºC (--●--), 20ºC (—●—) and 40ºC (–●–). Bars indicate standard deviation.

3.2 Imaging response

The response obtained from the imaging device evidenced the capacity of the technique to capture the disruption generated in the diffraction patterns due to the modification of the CO matrix. This first exploration of the obtained spectra showed differences in the laser profiles, mainly in the external zones, where the transmitted light was softer. Figures 4-A and 4-B can be visualised the effect of alterations in CO on the captured information by the imaging technique. Figure 4-A show the raw spectra from the RGB-lasers for pure CO and M, and the spectra obtained from the CO adulterated with 1% of M. In the same way, Figure 4-B compares pure CO from t0 of storage with CO stored 20ºC/63 days. Both examples showed alteration in the external zones of the profiles, however, with the aim of highlighting the differences, a mean-centering process for each example. Figures 4-C and 4-D show the mean-centered spectra, where the differences between the control CO and the altered CO samples can be observed in the above-mentioned zones of the profiles. The reduced modifications in the central zones would be related with the saturation of light. The high intensity of the laser-light transmitted in this zone (Airy disk) of the pattern produces a zone with a lower sensibility to the changes in the matrix than the extreme ones. Thus, after observing the capacity of the technique to capture the variability generated by the studied

alterations in the CO matrix, the possibilities of modelling those alterations from the imaging information were explored applying different machine learning procedures.

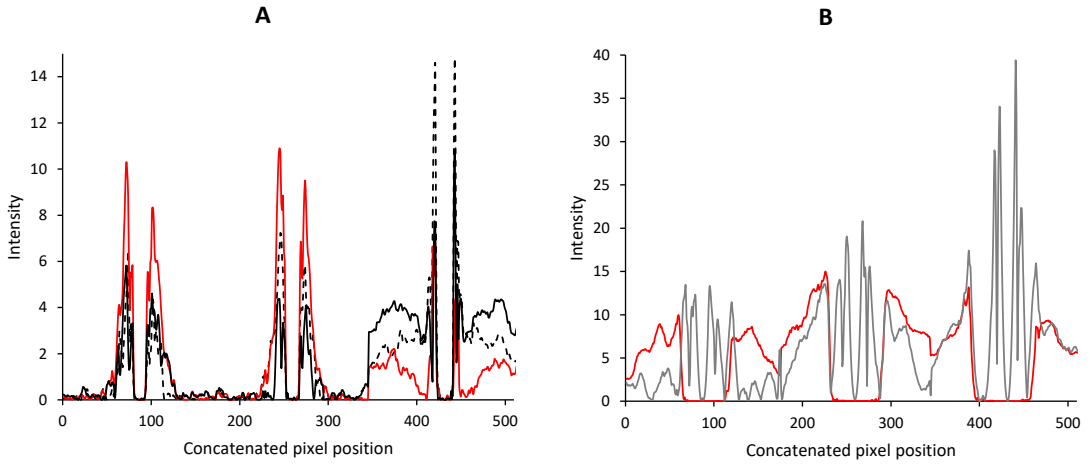

Figure 4. Examples of obtained RGB-lasers signal from different CO samples. A: mean-centered spectra from pure CO (—), pure M (—) and CO adulterated with M (5%) (- -); B: mean-centered spectra from pure CO and CO 20ºC/60days (—);

3.2 Modelling changes from adulterations (discrimination studies)

As it was above mentioned, the changes generated in CO due to adulterations were proposed in four phases focused to characterizing the pure adulterant oils and their presence in CO. The first test was focused to discriminate CO from the remaining oils. The results from PLS-DA models using imaging information from R, G, B and complete R+G+B spectra are included in the Table 2. The models performed from B-laser and complete R+G+B information had the highest significant F-scores, although R and G-models also provided successful results with F-scores > 0.95. These results evidenced the capture of the inner features of CO opposite to the properties of other similar oils, regardless of the type. The next model (OT) was focused on generating a model to not only discriminate CO but also recognize the other types of oils which could act as

adulterants. The results in Table 2 show the best F-score (0.98) for R+G+B information, followed by B-laser (0.93). In this case, increasing the oil categories increases the complexity the model, reducing the viability of the R and G-lasers.

After observing the capacity to discriminate CO and the adulterant oils, discrimination models of adulterated CO with those oils were tested. The first model in this sense was performed to discriminate pure CO from CO adulterated with any of the adulterants. The results are included in the Table 2 (A-CO). In this case, all of lasers reported F-scores>0.90, however the best classification was obtained from R+G+B complete information. The last model was focused on identifying the type of adulteration on the previous-identified adulterated CO samples. The results in the Table 2 (OT-A-CO) shows again the best results for R+G+B model. In this case F-score did not pass 0.90. That reduction manifests the increased difficulty in discriminating the type of adulterant after being mixed with CO in a reduced proportion (1 and 5%).

Table.2 Results of the modelling procedures

Model types		Laser light data as regressors			
Changes by adulteration (Discrimination models)		*R*	*G*	*B*	*RGB*
CO-M		0.97b	0.95a	0.99c	0.99c
OT		0.87b	0.78a	0.93c	0.98d
A-CO		0.92a	0.93b	0.94b	0.99c
OT-A-CO		0.77b	0.67a	0.86c	0.90c
Changes by oxidation (Regression models)					
Temperature of storage	Oxidation parameter				
4°C	*PV*	0.43a	0.44a	0.43a	0.92b
	AV	0.87b	0.62a	0.61a	0.89c
	TV	0.45a	0.46a	0.46a	0.89b
20°C	*PV*	0.90b	0.84a	0.89b	0.94c
	AV	0.87a	0.88a	0.86a	0.92b
	TV	0.79a	0.84a	0.74a	0.94b
40°C	*PV*	0.40a	0.40a	0.67b	0.91c
	AV	0.83a	0.80a	0.90b	0.91b
	TV	0.56a	0.54a	0.74b	0.91c
All samples	*PV*	0.46a	0.40a	0.61b	0.90c
	AV	0.69c	0.46b	0.31a	0.90d
	TV	0.49b	0.42a	0.66c	0.93d

Finally, after evidence the capacity of the technique to discriminate CO from the different types of oils and adulterated CO, a total discriminative model (T-CO) was performed including the oxidised samples detected in the analysis of the oxidation status (>10mEq O_2). Thus, the model was trained from pure CO and the rest of the non-pure CO samples as two categories. The F-scores were successful for all lasers (R=0.91, G=0.91 and B=0.92), however, in the same way of the previous models, complete R+G+B information rise the best result with a 0.96. Figure 5 shows the variance space within the two first LV from the PLS-DA, where the discrimination of CO samples can be visualised.

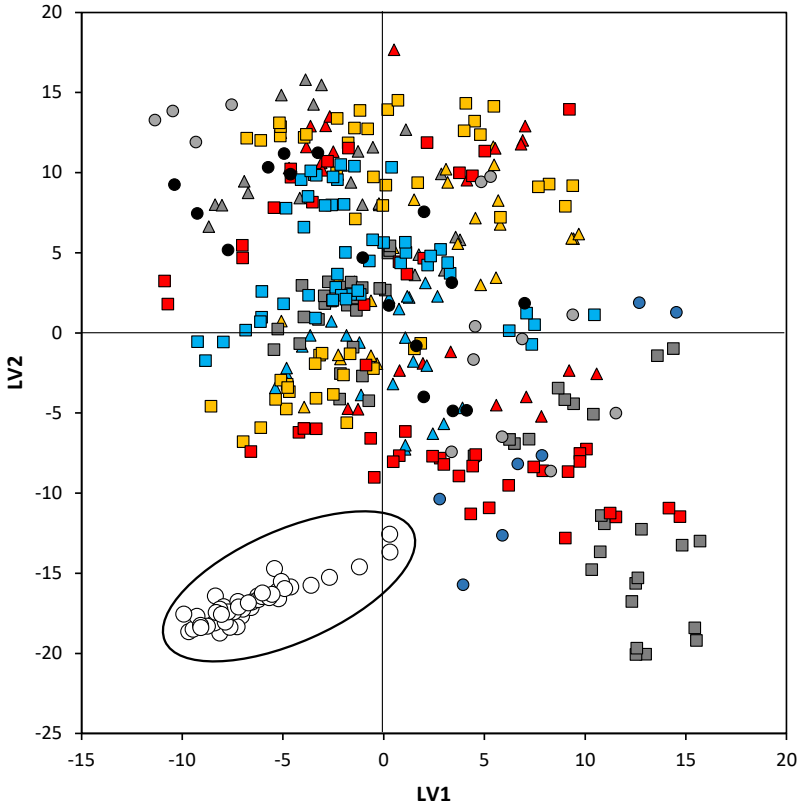

Figure 5. Representation of a part of the variance space (LV1 and LV2) from total CO discriminant model (T-CO) done with PLS-DA where CO samples can be distanced from the rest of the samples. ○:CO; ▲:WG; △:M; △:So; ▲:Se; ■:CO+WG; ■:CO+M; ■:CO+So; ■:CO+Se; ● :oxidised CO-4ºC; ● :oxidised CO-20ºC; ● :oxidised CO-40ºC.

The relationship between the imaging information in type of laser terms and the complexity of the model has been represented in the Figure 6. The number of latent variables (NLV) represent the complexity of the model. Models including two categories of samples (CO-M, A-CO and T-CO) raised F-scores>0.90 for all lasers, while the models where the adulterants were characterised (OT and OT-A-CO) only raised that mark with complete R+G+B information (and B in one case) at the same time that the cost of the model in NLV increased drastically. Thus, the changes from adulteration could be successfully modelled by discrimination procedures due to the high capacity of the technique to capture the variability among oils. The best results were obtained using the complete R+G+B information for most of the cases, principally for the models including more

types of oils. It meant that the captured diffraction patterns from the RGB lasers could fix information from which models with different complexity can be performed.

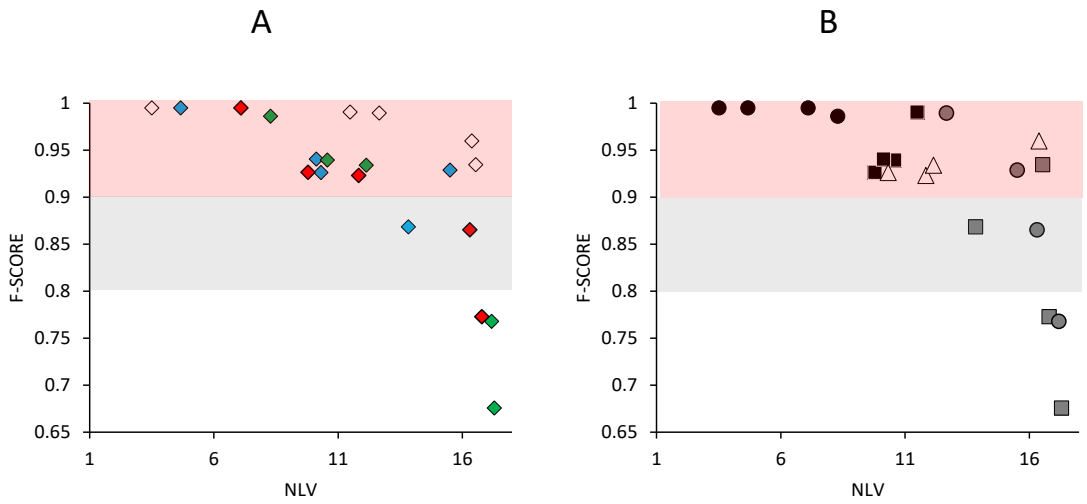

Figure 6. Discrimination model success and complexity in function of type of imaging information and samples. ◇: R+G+B models; ◆: R-laser models; ◆: B-laser models; ◆: G-laser models; ●:CO-M (2 categories); ●: OT (5 categories); ■:A-CO (2 categories); ■: OT-AC-O (5 categories);△: T- CO (2 categories). Red zone: F-score>0.9; grey zone: F-score 0.8-0.9.

The results reproduce those reported in other studies, where similar functions are developed based on other spectroscopic techniques. Abdul Rohman & Che Man (2011a) also reported models based on FTIR and PLS and PCR, where canola, corn, soybean and walnut oils were successfully detected at 1% (v/v) with a R^2>0.98 for all cases, but in that case NIR zone of the spectra was used. Moreover, Giese et al. (2019) also obtained successful results quantifying vegetable oil adulterations in CO with FT-IR spectra with a detection limit of 0.22%, and a 100% correct classification of pure cod liver using NMR.

3.3 Modelling changes from oxidation (regression studies)

The effect of the oxidation on CO during different storages was characterised by the imaging technique. The evolution of the chemically-measured parameters of oxidation status PV, AV and TV was modelled from the imaging information at each storage condition and including all samples as a common model. Table 2 shows the results in R^2-prediction terms. In this case, models performed from complete R+G+B information reported higher prediction capacity that the single lasers (~ 0.90). This result was observed for all storage temperatures. In regard to the oxidation parameter, there was not observed a pattern of better prediction for some of them. All had high and low predictions such the case of PV, which raised 0.90 with R-laser at 20ºC but 0.43 ate 4ºC with that same laser. When all samples were included in the model, similar results to the specific models were obtained. The R+G+B models provide R^2=0.9 for PV and AV and 0.93 for TV. Figure 7A and 7B shows the examples of observed vs predicted values for PV and AV. Results showed how incorporing more samples and image information generates more complex models (Figure 6). This tendency is clear in the evolution of the models including all samples. PV, AV and TV were successfully predicted from R+G+B information, however more than 8 NLV was the optimum, while models from the single temperatures required less than half NLV to obtain the same result (Figure 7C

). This pattern agreed with the previously observed in the classification models. These results evidenced that the variability generated in the imaging data by the oxidation processes had not dependence from the temperature of storage. It allowed to train a precise prediction model regardless the storage conditions of the selected samples.

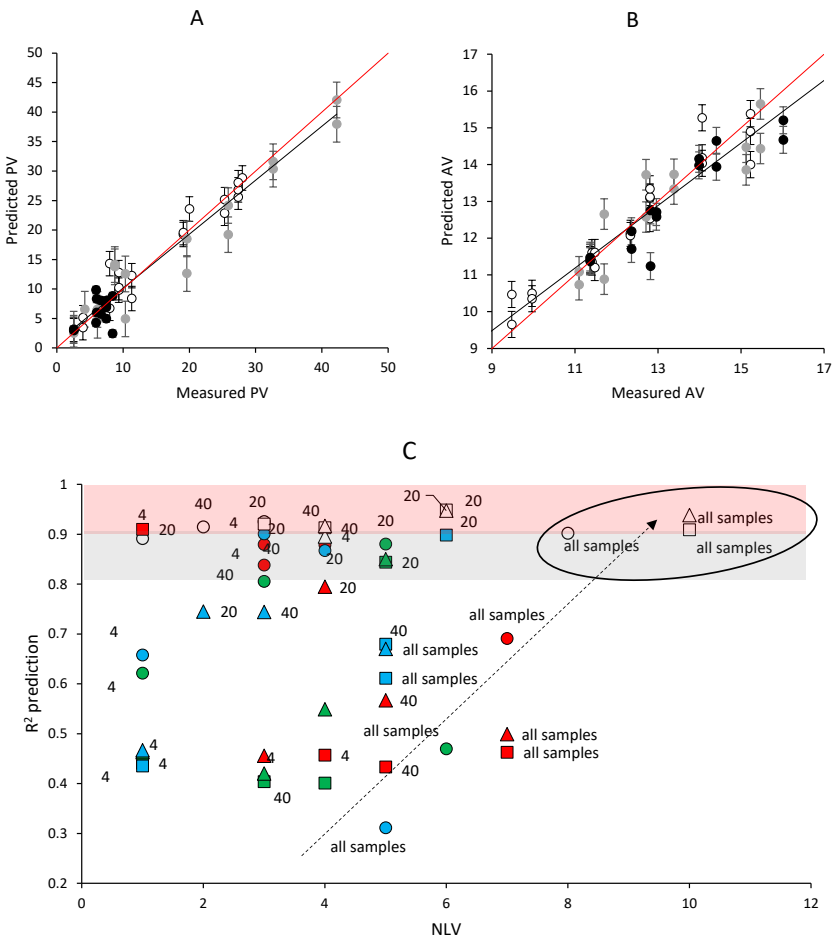

Figure 7. Measured and predicted PV (A) and AV (B). Red and black lines represent identity line 1:1 and real adjust, respectively. 4ºC (--●--), 20ºC (—●—) and 40ºC (–●–). Bars indicate standard deviation. Regression model success and complexity in function of type of imaging information and samples (C). ■: PV; ●: AV; ▲: TV; red: R-laser models; blue: B-laser models; green: G-laser models, white: R+G+B models. Red zone: R^2>0.9; grey zone: R^2 0.8-0.9. Arrow highlights the evolution of the models including all samples.

5. Conclusions

The RGB-laser scattering imaging technique combined with machine learning procedures was tested to characterise the alterations in the composition of cod-liver oil due to adulterations with other oils and oxidation. The results showed the capacity of the technique to capture the variability provided by the cod-liver oil against the other vegetable oils used as adulterants. The technique

discriminated the fish oil from pure wheat-germ oil, soybean oil, sesame oil and corn oil, at time to detect them when were added as adulterants. In the same way, changes in the oxidation status were successfully modelled by predicting the oxidation parameters with R^2>0.90, independently of the temperature of storage. Those capacities made possible to generate a discrimination model including all non-pure cod-liver oil samples, where pure cod-liver oil was differenced from other vegetable oils, adulterated cod-liver oil and non-eatable cod-liver oil due to oxidation status. The higher prediction capacity of the information captured by each type of laser was observed when complete spectra was used (R+G+B). The discriminative capacity was synergic when models were performed including the three lasers, as opposed to the results obtained with singly-laser models. In general, individual lasers did not generate accurate models, with some exceptions. Thus, the technique could be used as fast and non-destructive accessory to improve quality control procedures in a low-cost way, which can help to detect fraud in this type of fish oil. Following developments based on this technique must include more types of adulterants and contaminants from a non-oily origin, as well as and possible hazardous chemicals typical in sea products such as heavy metals and marine toxins.

6. Bibliography

Ahmmed, F., Killeen, D. P., Gordon, K. C., & Fraser-Miller, S. J. (2022). Rapid Quantitation of Adulterants in Premium Marine Oils by Raman and IR Spectroscopy: A Data Fusion Approach. *Molecules*. https://doi.org/10.3390/molecules27144534

AOCS. (2017). AOCS Official Method Cd 18-90: p-Anisidine Value. *Official Methods and Recommended Practices of AOCS*.

Cano Marchal, P., Martínez Gila, D., Gámez García, J., & Gómez Ortega, J. (2013). Expert system based on computer vision to estimate the content of impurities in olive oil samples. *Journal of Food Engineering*. https://doi.org/10.1016/j.jfoodeng.2013.05.032

Cortese, M., Riise, T., Bjørnevik, K., Holmøy, T., Kampman, M. T., Magalhaes, S., Pugliatti,

M., Wolfson, C., & Myhr, K. M. (2015). Timing of use of cod liver oil, a vitamin D source, and multiple sclerosis risk: The EnvIMS study. *Multiple Sclerosis*. https://doi.org/10.1177/1352458515578770

Geladi, P., & Kowalski, B. R. (1986). Partial least-squares regression: a tutorial. *Analytica Chimica Acta*. https://doi.org/10.1016/0003-2670(86)80028-9

Giese, E., Rohn, S., & Fritsche, J. (2019). Chemometric tools for the authentication of cod liver oil based on nuclear magnetic resonance and infrared spectroscopy data. *Analytical and Bioanalytical Chemistry*. https://doi.org/10.1007/s00216-019-02063-y

Lentjes, M. A. H., Welch, A. A., Mulligan, A. A., Luben, R. N., Wareham, N. J., & Khaw, K. T. (2014). Cod liver oil supplement consumption and health: Cross-sectional results from the EPIC-Norfolk cohort study. *Nutrients*. https://doi.org/10.3390/nu6104320

Rohman, A., & Che Man, Y. B. (2011). Analysis of chicken fat as adulterant in cod liver oil using Fourier transform infrared (FTIR) spectroscopy and chemometrics. *CYTA - Journal of Food*. https://doi.org/10.1080/19476337.2010.510211

Rohman, Abdul, & Che Man, Y. B. (2011a). Application of Fourier transform infrared (FT-IR) spectroscopy combined with chemometrics for authentication of cod-liver oil. *Vibrational Spectroscopy*. https://doi.org/10.1016/j.vibspec.2010.10.001

Rohman, Abdul, & Che Man, Y. B. (2011b). Authentication analysis of cod liver oil from beef fat using fatty acid composition and FTIR spectra. *Food Additives and Contaminants - Part A*. https://doi.org/10.1080/19440049.2011.600727

Romano, G., Argyropoulos, D., Nagle, M., Khan, M. T., & Müller, J. (2012). Combination of digital images and laser light to predict moisture content and color of bell pepper simultaneously during drying. *Journal of Food Engineering*, *109*(3), 438–448. https://doi.org/10.1016/j.jfoodeng.2011.10.037

Sacks, F. M., Lichtenstein, A. H., Wu, J. H. Y., Appel, L. J., Creager, M. A., Kris-Etherton, P.

M., Miller, M., Rimm, E. B., Rudel, L. L., Robinson, J. G., Stone, N. J., & Van Horn, L. V. (2017). Dietary fats and cardiovascular disease: A presidential advisory from the American Heart Association. In *Circulation*. https://doi.org/10.1161/CIR.0000000000000510

Verdú, S., Barat, J. M., & Grau, R. (2019). Laser backscattering imaging as a non-destructive quality control technique for solid food matrices: Modelling the fibre enrichment effects on the physico-chemical and sensory properties of biscuits. *Food Control*. https://doi.org/10.1016/j.foodcont.2019.02.004

Verdú, S., Perez, A. J., Barat, J. M., & Grau, R. (2021). Laser-backscattering imaging for characterising the dairy matrix in different phases during curd processing. *Food Control*. https://doi.org/10.1016/j.foodcont.2021.108193

Verdú, S., Pérez, A. J., Barat, J. M., & Grau, R. (2018). Laser backscattering imaging as a control technique for fluid foods: Application to vegetable-based creams processing. *Journal of Food Engineering*, *241*(May 2018), 58–66. https://doi.org/10.1016/j.jfoodeng.2018.08.003

Verdú, S., Ruiz-Rico, M., Pérez, A. J., Barat, J. M., & Grau, R. (2020). Application of laser backscattering imaging for the physico-chemical characterisation of antimicrobial silica particles functionalised with plant essential oils. *Journal of Food Engineering*. https://doi.org/10.1016/j.jfoodeng.2020.109990

Xie, C., Wang, Q., & He, Y. (2014). Identification of different varieties of sesame oil using near-infrared hyperspectral imaging and chemometrics algorithms. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0098522

Joint FAO/WHO Codex Alimentarius Commission, 2017. Codex Alimentarius: Standard for fish oils CXS 329–2017. Rome: World Health Organization. Food and Agriculture Organization of the United Nation (2017)