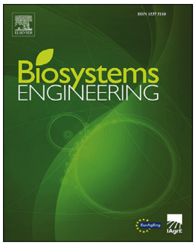
Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110



Research Paper

Use of visible and near infrared spectroscopy with a view to on-line evaluation of oil content during olive processing

Valentina Giovenzana [a], Roberto Beghi [a,*], Roberto Romaniello [b], Antonia Tamborrino [c], Riccardo Guidetti [a], Alessandro Leone [b]

[a] *Department of Agricultural and Environmental Sciences — Production, Landscape, Agroenergy, Università degli Studi di Milano, via Celoria 2, 20133 Milano, Italy*
[b] *Department of the Science of Agriculture, Food and Environment, University of Foggia, via Napoli 25, 71100 Foggia, Italy*
[c] *Department of Agricultural and Environmental Science, University of Bari Aldo Moro, via Amendola 165/A, 70126 Bari, Italy*

ARTICLE INFO

Article history:
Received 8 May 2017
Received in revised form
7 May 2018
Accepted 9 June 2018

Keywords:
By-products
Chemometrics
Extra virgin olive oil
Extraction process
Optical analysis
Predictive model

The aim of this preliminary feasibility study was to verify whether visible/near infrared(vis/NIR) spectroscopy could be used to predict the oil content of intact olives entering the mill, and of olive paste, pomace and paté during the milling process. Three different extraction methods (3-phase decanters, 2-phase decanters and 2.5-phase decanters) were considered, and two optical devices were tested: (i) a process device for non-contact analysis and (ii) a system equipped with an immersion probe for contact measurements, both working in the spectral range 400–1650 nm. 35 samples of olives were collected during the experimental tests, 50 samples of olive paste, 50 samples of pomace and 50 samples of paté. The collected samples (olives, olive paste, pomace and paté) were used to calculate partial least squares (PLS) regression models.

Results regarding the non-contact analyses were encouraging, except for the measures on olives. On pomace, satisfactory models were calculated for the vis/NIR range [Ratio Performance Deviation (RPD) > 2], and a good model with $R^2 = 0.81$ and RPD = 2.68 in validation was calibrated in the NIR range. The device equipped with an immersion probe achieved good predictive models for the oil content prediction on paté (R^2 and RPD values ranged 0.77–0.82 and 3.00–3.43).

The predictive models could be easily applied in an on-line system to monitoring the entire extraction plant and to perform a feed-forward control, allowing a reduction of oil leakage to minimise the oil losses and to maximise the extraction yield.

---

* *Corresponding author*. Università degli Studi di Milano, via Celoria 2, Milano 20133, Italy.
E-mail address: roberto.beghi@unimi.it (R. Beghi).

https://doi.org/10.1016/j.biosystemseng.2018.06.001

1. Introduction

The food industry converts agro-materials into products using a variety of different processes. The choice of the process to achieve the best supply chain performance is carried out according to the efficiency and quality of the product to remain competitive on the market. For the production of an extra-virgin olive oil of excellent quality, a company should consider implementing and/or improving some parts of the olive oil production chain: olive harvesting and handling; carrying out the milling operation a short time after olive harvesting; using advanced mill plant equipped with suitable technologies to control the milling conditions. The last point is a good starting point for an optimisation of the process because differences from the standard operating conditions are the principal cause of failures to maintain the highest standards during processing and these affect the final product's quality. In the extra virgin olive oil extraction chain, control of the process and its management determines the conditions to produce a high quality extra-virgin olive oil, needed to retain consumer confidence, but process control is also important to evaluate potential losses of the plant's yield. Thus, process control is essential for achieving effective competitiveness.

The extra virgin olive oil extraction plant is set up as a number of interconnected machines arranged in series and in some cases, even in parallel (Leone et al., 2015). The main operations of the entire process are: olive cleaning, olive crushing to obtain a paste, paste malaxation, solid–liquid separation and liquids separation. Solid-liquid separation is one of the most important operations; it consists of separating the solids (pomace) from the other olive components (oil and waste water).

The solid–liquid centrifuges are classified according to the characteristics of the products in output from the decanter (Altieri, Di Renzo, & Genovese, 2013; Leone, Romaniello, Zagaria, & Tamborrino, 2015):

- 3-phase decanters, able to separate oil (5–10% of humidity) from dry pomace (50–55% of humidity) and waste water. In three-phase sedimentation, 10–30% of warm water is added to the inlet olive paste;
- 2-phase decanters, able to separate oil (5–10% of humidity) from the wet pomace (65–70% humidity). In two-phase sedimentation, no water is added to the inlet olive paste;
- 2.5-phase decanters, able to separate oil (5–10% humidity) from dry pomace rich in heavy solids at 50–55% humidity and to produce a new by-product called "paté," which contains wastewater enriched by soft solids without pit fragments and having a semi-solid consistency.

An important evaluation parameter for a decanter's performance is the oil extractability. This parameter is calculated by dividing the oil extracted by 100 kg and the oil contained in those 100 kg of olives. The decanter's extractability can vary according to (i) the type of decanter, (ii) the wear/maintenance, (iii) the cultivar and ripeness of the olives processed, (iv) the particle size and rheology of olive paste and finally on (v) the decanter adjustments. Extractability is in the range 80–90% (Tamborrino, Leone, Romaniello, Catalano, & Bianchi, 2015); the oil that is not extracted remains in the unbroken cells or is trapped in the cytoplasm tissues, or is emulsified into the aqueous phase. The extractability reflects the decanter's correct functioning, and it is very important to know it. The quantification of residual oil in the pomace is considered a crucial control for the qualitative optimisation of the olive oil extraction plant.

To obtain on-line information on the oil content in input olives and in the output by-products allows managers to react to imperfections in the process by corrective actions or by redefining/reinforcing preventive actions. Nowadays the traditional or rapid Soxhlet method is used to analyse the oil content in olives, pomace and paté, requiring a time-consuming drying step, followed by an extraction using solvent. This method is now often substituted in routine analyses by nuclear magnetic resonance (NMR) spectroscopy, but this technique is also not very rapid due to water interference. For this reason, the olive pomace sample must be completely dry. Consequently, this method is unsuitable for an on-line application during process control.

The investigation of the characteristics of the olives entering the milling process and of the features of by-products during milling could allow operators to control the quality of the process. A better monitoring of the oil production process also depends on controlling production of the paste, the pomace and the paté, the intermediate products between the olives entering into the process and the oil outlet from the mill, to establish correlations among olives, paste, pomace, paté and oil. Hence, in the olive oil industry, quick and easy-to-use technologies are required to (i) assess olive ripening and the characteristics of the by-products, (ii) for early detection of possible failures, (iii) to monitor in a lasting way the production process during its crucial steps in order to control the oil quality and yield. The sector could be helped by optical non-destructive and rapid applications for the olive oil production chain optimisation.

Over the last 30 years, on/in-line NIR spectroscopy has gained much success by placing on the market efficient and advanced tools for continuous product quality monitoring in the food processing industry, such as for fruit, vegetables, meat, grain, dairy products and beverages (Huang, Yu, Xu, & Ying, 2008; Porep, Kammerer, & Carle, 2015). Regarding the oil sector, several studies have highlighted the enormous opportunities offered by NIR spectroscopy in terms of applications for quality control during the process, performing on/in/at-line measurements on olive fruits, on pastes, and on oils (Armenta, Moros, Garrigues, & Guardia, 2010). Researchers tend to focus attention on the on-line applications of non-invasive technologies in order to reduce the gap between laboratory scale experimentation and the olive milling industry (Ortega, Gila, Puerto, García, & Ortega, 2016). A number of studies applying different vibrational techniques in the olive oil chain can be found in the literature, mainly with the aim of standardising the procedure for an application as official control of the end product (Nenadis & Tsimidou, 2017). For this purpose, it is crucial to evaluate the optimal spectral range to be used and the chemometric methods to be performed to obtain robust predictive models for the estimated parameters. On intact olives, Giovenzana, Beghi, Civelli, Marai, and Guidetti (2015) and Beghi, Giovenzana, Civelli, Cini, and Guidetti (2013) studied the capability of portable visible/near infrared (vis/NIR) and NIR spectrophotometers to investigate different texture indices for the characterisation of olive fruits entering the milling process. Salguero-Chaparro, Baeten, Fernández-Pierna, and Peña-Rodríguez (2013) used NIR spectroscopy on intact olives for the on-line determination of the oil content, moisture and free acidity parameters. NIR has been used in the laboratory or in processing mill lines for the analysis of olive by-products such as olive pomace. On olive paste and pomace, García Sánchez, Ramos Martos, and Ballesteros (2005) tested the suitability of NIR and NMR spectroscopy for the determination of moisture and fat contents, while Gallardo, Osorio, and Sanchez (2005) used near-infrared spectroscopy for the real-time determination of moisture and fat content in olive pastes and pomace obtained from 2-phase and 3-phase decanters; Barros, Nunes, Martins, and Delgadillo (2009) applied Fourier transform Near Infrared Spectroscopy (FTNIR) spectrometry (1000–2500 nm) in combination with partial least squares regression for direct, reagent free determination of fat and moisture content in milled olives and olive pomace; Muik, Lendl, Molina-Díaz, and Pérez-Villarejo (2004) compared NIR reflectance and Raman spectrometry to determine the oil and water content of olive pomace. Marquez, Díaz, and Reguera (2005) used an optical NIR sensor for on-line characterisation of virgin olive oil.

The aim of this preliminary study is to verify the applicability of two different vis/NIR spectroscopy devices to predict the oil content of intact olives entering the mill and then to evaluate the oil content in olive paste, pomace and paté during the milling process. Three different extraction methods will be considered. The predictive models determined starting from vis/NIR spectra on olive pastes and by-products during the production process could be applied in an on-line system to monitor the entire extraction plant and to perform a feed-forward control to minimise oil losses and to maximise the extraction yield.

2. Material and methods

2.1. *Industrial olive oil extraction plants and samples collection*

Experimental tests were performed in two different full-scale mills during the cropping season 2015–2016. The first plant was located in Foggia and the second one in Trani (BT).

The first mill was constituted by a defoliator, a washing machine, a partial de-stoning machine (Pietro Leone e Figli s.n.c., Foggia - Italy), a series of six malaxer machines, a 3-phase solid–liquid horizontal centrifugal decanter (mod. NX X32, Alfa Laval Corporate AB, Lund, Sweden), and two liquid–liquid vertical plate centrifuges. The malaxer is a machine in which the olive paste is kneaded and thermally conditioned. The 3-phase decanter discharges 2 liquid phases (waste water and oil) and one solid phase (dry pomace). Mass flow rate of the plant was 3000 kg h^{-1} and the malaxation was performed for 30 min at 27 °C (Fig. 1).

The second mill was equipped with a hammer crusher (model Frangolea, Barracane s.r.l., Modugno, BA, Italy), a set of four malaxer machines (model Gramola 3000, Barracane s.r.l., Modugno, BA, Italy), connected in series, a decanter centrifuge model Megala 650 CI (Barracane s.r.l., Modugno, BA, Italy) and a liquid–liquid vertical plate centrifuge (model Grande 3000, Barracane s.r.l., Modugno, BA, Italy). A Megala 650 CI decanter centrifuge can operate in 2.5-phase, which produces dry pomace rich in rigid solids and paté. During the experimental test the mass flow rate was set to 6500 kg h^{-1} and the malaxation was performed for 50 min at 27 °C.

Olives, olive pastes and pomace in the Foggia mill and olives, pomace and paté in the Trani mill, were collected at regular time intervals.

35 samples of olives, 50 samples of olive paste, 50 samples of pomace and 50 samples of paté were collected during the experimental tests.

All samples were subjected to non-destructive analysis, acquiring the vis/NIR spectra immediately after the sampling in an experimental station installed in the olive mill. Subsequently, the acquired samples were sealed and stored for 4 days in the cooling room at 2 °C before the laboratory analyses.

All tests were conducted on olive fruits of the ‘Coratina’ cultivar (*Olea europaea* L.) with a maturity index value in the range 1.8–2.0. The fruit's ripeness was determined according to the proposed method of the International Olive Council (IOOC, 2001). The olives were purchased by the mills' owners. The olives were harvested a few hours before the experimental tests, near the mills.

2.2. *Visible/near infrared systems*

The spectral acquisitions were performed using two optical devices on samples taken from the different sections of the

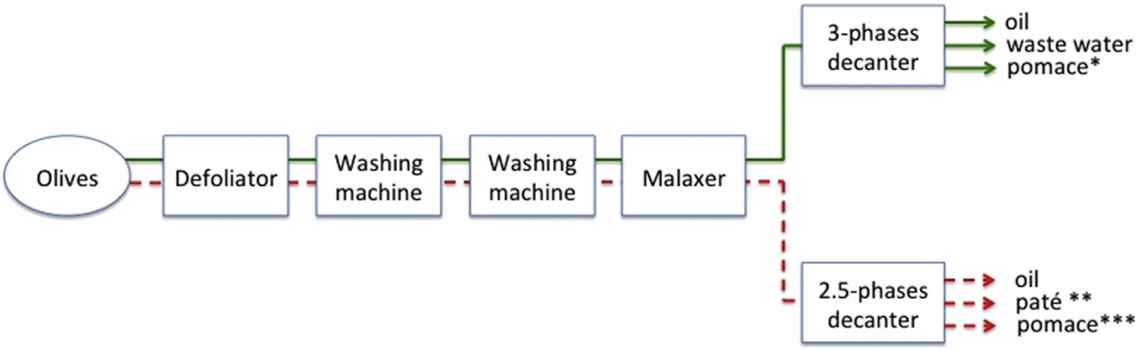


Fig. 1 – Extraction lines considered for the experiments (solid line = extraction by 3-phase decanter; dash line = extraction by 2.5 phase decanter). * soft solids (pulp + skin); ** wastewater + soft solids fraction; * dark solids + soft solids fractions.**

plants, simulating the measurement conditions foreseeable when using the devices installed on the plants: (i) a device for non-contact analysis at a variable distance between sensor and sample and (ii) a system equipped with an immersion probe for contact measurements.

2.2.1. *Device I*

The samples' spectral acquisitions were performed using a process vis/NIR device (Corona Process, Zeiss, Germany) built to work in operative process conditions for non-contact analysis of flows and/or non-homogeneous products (device features: dimension 400 (w) × 300 (l) × 300 (p) mm, weight about 15 kg). The system carries out measurements in the vis/NIR spectral range (380–1650 nm, spectral resolution 2.0 nm) at a variable autofocused distance between sensor and sample in the range 80–600 mm. The device is therefore ideally suited for on-line application during the milling process. The system will allow the operators to define real-time objective indicators on the product usefully for better plant management.

The optical device consisted of four components: 1) a lighting system (halogen lamp) with auto calibration of white and black references, 2) a spectrophotometer, 3) hardware for data acquisition and instrument control, and 4) the power supply.

Spectra were acquired in reflectance mode: the sample's surface was irradiated by the light source through a 60 mm spotlight. Each spectral sample was obtained by averaging three acquisitions in three different spots of the batch. Each acquisition represents an average of ten reflectance spectra. The measurements were performed at about 400 mm between sensor and probe (olives, olive paste and pomace).

2.2.2. *Device II*

The optical analysis was also carried out using a laboratory spectrophotometer MCS 600 (Zeiss, Germany) characterised by a double acquisition module: 450–950 nm and 950–1650 nm, spectral resolution 3.0 nm. The instrument was equipped with a probe for contact and immersion reflection measurements (particularly suited for non-solid samples such as paté). In the optic probe the fibres are arranged in a 6-around-1 configuration. Spectra were acquired through a Y-shaped, bidirectional fibre optic probe. A Y-shaped fibre allowed the light coming from the halogen lamp to be guided to illuminate the sample while simultaneously collecting the radiation coming from it and guiding it back to the spectrophotometer. The internal components of Device II are similar to those explained for Device I.

The measurements on olive paste, on olive pomace and on paté samples were performed using Device II.

2.3. *Oil content in olives, paste, pomace and paté*

The total oil mass fraction was determined on a 40 g sample that was previously dehydrated until it reached a constant weight. The olive samples were also previously milled. The oil was extracted by hexane using an automatic extractor (Randall 148, Velp Scientifica, Milano, Italy) following the analytical technique described by Cherubini et al. (2009).

2.4. *Statistical analysis*

The acquired data were processed using chemometric techniques to extract maximum data information. Chemometric analysis was performed using The Unscrambler software package (version 9.8, CAMO ASA, Oslo, Norway). Four pre-treatments were applied to the vis/NIR and NIR spectra to maximise the model accuracy. Moving-averaged smoothed spectra (15 point-wide window corresponding to a window of 4.5), multiplicative scatter correction (MSC), standard normal variate (SNV) and Savitzky Golay first derivative (der1) were applied before building the calibration models. Pre-treatments were applied to improve the signal-to-noise ratio in order to reduce spectral noise. The available samples were used to calculate the chemometric regression model for reference parameters using partial least squares (PLS) regression analysis. The PLS method is particularly suitable in the case of high correlation among variables, i.e. for spectral data of intact fruit and vegetables (Nicolaï et al., 2007).

A procedure for outlier detection during model calibration was performed. The vis/NIR and NIR spectra were checked visually for abnormal spectra, which may result from incorrect sampling (Lachenmeier, 2007). Moreover, the “influence plot” function of The Unscrambler software was studied for the detection of outliers. For the determination of the number of samples considered as outliers not to be included in model calibration, the “Hotelling T2 Computations” function was applied (α value 0.05). To evaluate model accuracy the following were considered: the coefficient of determination in calibration (R^2_{cal}), the coefficient of determination in cross-validation (R^2_{cv}), root mean square error of calibration (RMSEC), and root mean square error of cross-validation (RMSECV). Calibration models were evaluated using a cross-validation leave-more-out procedure using five groups for the training set (cancellation groups), randomly selected. The use of a small number of cancellation groups contributes to avoiding model overfitting (Casale, Casolino, Ferrari, & Forina, 2008). Due to different sample sizes used for each model, the sample number incorporated for each cancellation group ranged from 5 to 18 samples. Moreover, the Ratio Performance Deviation (RPD) value was calculated. RPD is defined as the ratio between the standard deviation of the response variable and RMSECV. RPD values below 1.5 indicate that the calibration is not useful. When the RPD value is higher than 2, quantitative predictions are possible. Values were in the range 1.5–2.0, indicating that the algorithm has the possibility to distinguish between high and low values (Williams & Norris, 2001). The best model calibrations were selected based on minimising the RMSECV and maximising R^2 and RPD.

3. Results and discussion

Figure 2 shows the spectra acquired on intact olive, olive paste and pomace using Device I (Fig. 2a), and for olive paste, pomace and paté, using Device II (Fig. 2b). A visual analysis of the spectra highlights the principal wavelength peaks and trends allowing a clear separation of spectra from intact olives, olive paste, pomace and paté. The comparison of the

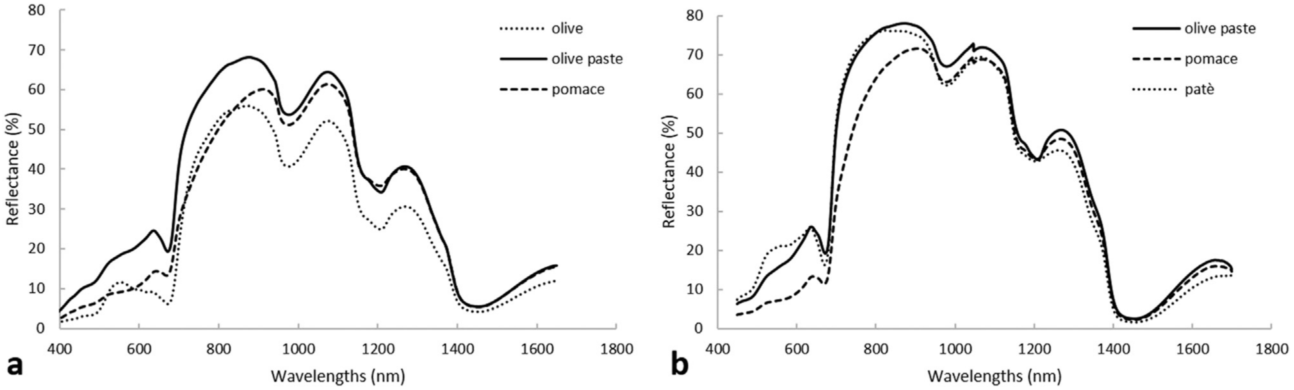


Fig. 2 – Spectra coloured based on the different analysed matrices using device I (a) and device II (b).

three spectral groups shows wide differences in the 550–680 nm range, due to external colour differences in the samples which are related to changes in the amount of pigment (related to the presence of anthocyanins) causing a decrease in reflectance in the visible band associated with the absorption peak at 540 nm. A high reflectance absorption could be noticed around the chlorophyll absorption peak at 680 nm (McGlone, Jordan, & Martinsen, 2002).

In the NIR region, a reflectance peak is noticeable around 970 nm (absorption peak relative to the second overtone of the water O–H bond in the near infrared region). The in-depth peak at 970 nm is characteristic of the vis/NIR measurements on products very rich in water, like fruits. Also, the peaks around 1180 nm and 1440 nm arise from typical water absorption (Kavdir, Buyukcan, Lu, Kocabiyik, & Seker, 2009; Williams & Norris, 2001).

Results for the PLS regression models from spectra acquired using Device I for the prediction of the oil content starting from different sample matrices collected in correspondence with critical steps of the milling process (olives entering the mill, paste during the malaxation, pomace and paté in the outlet of the decanter) are shown in Table 1. A dataset combining olive paste and pomace spectra was also used for the model calculation. Since (i) absorption of OH bonds relating to water content and, consequently, to oil content, were found in the NIR region and (ii) the acquisition in the visible region was affected by the olive surface and pulp colour increase during ripening causing a disturbance in the oil content estimation, specific models deriving only from the NIR spectral range (700–1650 nm) were also calculated. Overall, the results of this preliminary study are encouraging, except for measures on olives. In this preliminary study, the data set available was not particularly wide and the variability of the sample was not a-priori controlled but was due to the natural variability found in samples at the full-scale plants. This was probably due to the use of olives from the same cultivar, with a ripeness index in a small range 1.8–2.0 and harvested from the same production area. Models performance could be improved once calibrated under operating conditions using larger sample datasets from self-learning calibration systems and mostly using olives from different cultivars and with a wide range of ripeness index, for example between 1 and 3.5.

Good results were obtained using the wider combined dataset with the paste and pomace samples, with R^2 and RPD higher than 0.90 and 3.30 respectively, for both the measured spectral ranges. In general, slightly better results were obtained using the NIR range. On pomace, satisfactory models were calculated for the vis/NIR range 380–1650 nm (RPD > 2), and a good model with $R^2 = 0.81$ and RPD = 2.68 in validation was calibrated in the NIR range (700–1650 nm) using SNV and der1 pretreatments. Models developed for intact olives and olive paste did not reach sufficient predictive performance considering R^2 and RPD values in validation. For intact olives was not possible to calibrate a model for the NIR range due to the very low correlation between spectra acquired from a distance of 400 mm and olives' oil content in this range. Slightly better results were also obtained using the visible spectral range. This is probably due to the very low number of samples in this preliminary study and to the low correlation between the unevenness of the external characteristics of olive samples and the internal oil content. Some studies regarding the application of vis/NIR spectroscopy to intact olives can be found in the literature, reporting good results with R^2 ranged 0.70–0.90 (Stella et al., 2015). The authors, in a previous feasibility study, used a portable vis/NIR spectrophotometer for the estimation of the texture of the olives entering the mill in a lab-scale plant, with encouraging results ($R^2 = 0.68$ and RMSECV% = 8.2%) (Beghi, Giovenzana, Civelli, Cini, & Guidetti, 2013). However, most of the published studies concern the monitoring of ripening evolution, exploiting the advantages for the calibration of the models from the high variability of the data which are also due to individual olive's surface colour changes. Moreover, some studies were performed in more controlled sampling conditions where measurements were carried out through direct contact between the sensor and the single berry (Bellincontro, Caruso, Mencarelli, & Gucci, 2013).

Table 2 shows the results for the PLS regression models from spectra acquired using the Device II for the prediction of the oil content starting from different sample matrices collected in different points of the milling process (paste during the malaxation, pomace separation at the 3-phase decanter, paté separation at the 2.5-phase decanter). Also in this case a dataset combining olive paste and pomace spectra was used for the model calculation, and specific models deriving from the NIR spectral range (700–1650 nm) only were calculated. Device II was particularly suitable for liquid samples thanks to the immersion probe. For this reason, good

Table 1 – Descriptive statistics and statistics of the PLS models elaborated on vis/NIR and NIR spectra acquired by Device I, for the oil content estimation (g).

Matrix	Spectral Range (nm)	Pretreatments	n° samples	Mean	SD	Calibration				Validation			
						CV%	LV	R^2_{cal}	RMSE	R^2_{val}	RMSECV	RMSECV%	RPD
Olives	380–1650	smooth, MSC	27	6.11	0.93	15.2	3	0.40	0.51	0.30	0.56	9.17	1.67
		smooth, SNV, der1	27	6.11	0.93	15.2	3	0.36	0.49	0.22	0.55	9.00	1.70
	700–1650	smooth, MSC	27	6.11	0.93	15.2	2	0.15	0.74	0.08	0.82	13.42	1.13
		smooth,SNV, der1	27	6.11	0.93	15.2	2	0.12	0.86	0.10	0.90	14.72	1.03
Olive paste	380–1650	smooth, MSC	44	6.92	1.62	23.4	11	0.66	0.82	0.31	1.18	17.05	1.37
		smooth, SNV, der1	44	6.92	1.62	23.4	7	0.74	0.78	0.59	1.03	14.88	1.57
	700–1650	smooth, MSC	44	6.92	1.62	23.4	6	0.60	0.92	0.38	1.16	16.76	1.40
		smooth, SNV, der1	44	6.92	1.62	23.4	7	0.67	0.82	0.48	1.06	14.88	1.57
Pomace	380–1650	smooth, MSC	37	1.63	0.27	16.6	8	0.89	0.08	0.76	0.13	7.98	2.06
		smooth, SNV, der1	37	1.63	0.27	16.6	11	0.95	0.06	0.77	0.13	7.98	2.06
	700–1650	smooth, MSC	37	1.63	0.27	16.6	8	0.74	0.12	0.41	0.18	11.04	1.49
		smooth, SNV, der1	37	1.63	0.27	16.6	7	0.90	0.07	0.81	0.10	6.13	2.68
Olive paste + pomace	380–1650	smooth, MSC	90	4.07	2.86	70.3	3	0.91	0.81	0.91	0.86	21.13	3.32
		smooth, SNV, der1	90	4.07	2.86	70.3	4	0.93	0.75	0.91	0.86	21.13	3.32
	700–1650	smooth, MSC	90	4.07	2.86	70.3	3	0.92	0.80	0.91	0.84	20.64	3.40
		smooth, SNV, der1	90	4.07	2.86	70.3	4	0.93	0.73	0.91	0.82	20.15	3.49

PLS = Partial Least Square; smooth = smoothing; MSC = multiplicative scatter correction; SNV = standard normal variate; der1 = first derivative; SD = standard deviation; CV% = coefficient of variation (%); LV = latent variables; RMSE = root mean square error; RMSECV = root mean square error of cross validation; RPD = Relative percent deviation.

Table 2 – Descriptive statistics and statistics of the PLS models elaborated on vis/NIR and NIR spectra acquired by Device II, for the oil content estimation (g).

Matrix	Spectral Range (nm)	Pretreatments	n° samples	Mean	SD	Calibration				Validation			
						CV%	LV	R^2_{cal}	RMSE	R^2_{val}	RMSECV	RMSECV%	RPD
Olive paste	450–1650	smooth, MSC	42	6.92	1.62	23.4	2	0.30	0.85	0.25	0.89	12.86	1.82
		smooth, SNV, der1	42	6.92	1.62	23.4	3	0.54	0.69	0.44	0.77	11.13	2.10
	700–1650	smooth, MSC	42	6.92	1.62	23.4	3	0.35	0.85	0.12	1.03	14.88	1.57
		smooth, SNV, der1	42	6.92	1.62	23.4	4	0.56	0.75	0.48	0.84	12.14	1.92
Pomace	450–1650	smooth, MSC	36	1.63	0.27	16.6	2	0.29	0.18	0.24	0.19	11.66	1.41
		smooth, SNV, der1	36	1.63	0.27	16.6	4	0.52	0.17	0.39	0.19	11.66	1.41
	700–1650	smooth, MSC	36	1.63	0.27	16.6	2	0.41	0.16	0.24	0.20	12.27	1.34
		smooth, SNV, der1	36	1.63	0.27	16.6	5	0.59	0.13	0.32	0.17	10.43	1.59
Paté	450–1650	smooth, MSC	44	0.72	0.48	66.7	5	0.84	0.12	0.79	0.14	19.44	3.43
		smooth, SNV, der1	44	0.72	0.48	66.7	4	0.83	0.12	0.77	0.14	19.44	3.43
	700–1650	smooth, MSC	44	0.72	0.48	66.7	4	0.86	0.12	0.82	0.14	19.44	3.43
		smooth, SNV, der1	44	0.72	0.48	66.7	3	0.81	0.14	0.77	0.16	22.22	3.00
Olive paste + pomace	450–1650	smooth, MSC	88	4.07	2.86	70.3	5	0.93	0.74	0.91	0.83	20.39	3.44
		smooth, SNV, der1	88	4.07	2.86	70.3	4	0.94	0.68	0.92	0.75	18.43	3.81
	700–1650	smooth, MSC	88	4.07	2.86	70.3	4	0.91	0.85	0.89	0.95	23.34	3.01
		smooth, SNV, der1	88	4.07	2.86	70.3	3	0.92	0.78	0.92	0.82	20.15	3.49

PLS = Partial Least Square; smooth = smoothing; MSC = multiplicative scatter correction; SNV = standard normal variate; der1 = first derivative; SD = standard deviation; CV% = coefficient of variation (%); LV = latent variables; RMSE = root mean square error; RMSECV = root mean square error of cross validation; RPD = Relative percent deviation.

predictive models were calculated for the oil content prediction on paté (R^2 and RPD values ranged 0.77–0.82 and 3.00–3.43). MSC pretreatment on smoothed spectra showed better results for the NIR range.

Good results were also obtained using the combined dataset with the paste and pomace samples, for both the vis/NIR and NIR spectral ranges. R^2 and RPD values ranged 0.89–0.92 and 3.01–3.81, respectively. Der1 pretreatment gave better model performances for both the spectral ranges. Figure 3a shows an example of a PLS model calculated for the prediction of oil content (a) in pomace samples using device I, and SNV and der1 as pre-treatments; Fig. 3b shows the model for paté samples, using device II and MSC as pre-treatment. Bendini et al. (2007) applied FT-NIR for in-process monitoring of different cultivar pastes in diffuse reflectance mode, obtaining models with R^2 equal to 0.92 for the prediction of oil content.

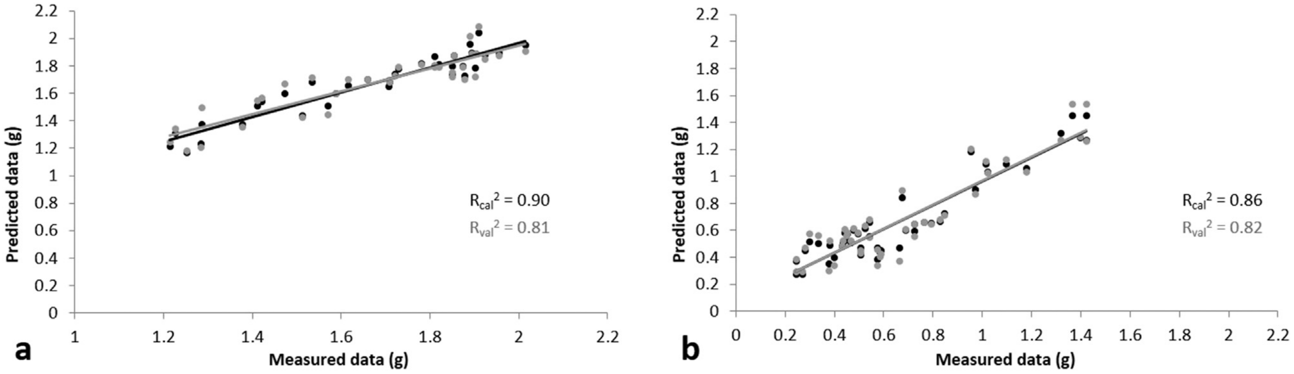


Fig. 3 – Examples of PLS model calculated for the prediction of oil content (g) in (a) pomace, using device I and (b) paté, using device II.

Predictive models obtained for the paté were promising and usable for the quantification of the residual oil content in a view of a real-time control and adjustment of the extraction process using the 2.5-phases decanter. Moreover, predictive models obtained for the paste and pomace dataset using both the devices are usable for the monitoring of operative parameters in different steps of the milling process, for the enhancement of extraction oil yield and the control of semi-finished products of the process. Studies reported in the literature concerning the application of NIR spectroscopy on olives for the estimation of the oil content reported results with $R^2 > 0.81$ (Ortega, Gila, Puerto, García, & Ortega, 2016). Most of these studies, however, involved experiments conducted on lab-scale or pilot-scale plants. No literature is available regarding the vis/NIR and NIR applications on semi-finished products of the process (i.e. paste, pomace and paté) sampled directly in full-scale plants. In fact, a requirement of efficient food processing is to ensure that all the product is obtained according to the production protocols and specially to identify in real-time the performance of single process steps and single machines. This would allow operators to identify in real-time the batches that did not meet agreed standards and the machines that do not give the best results. This reactive approach allows managers to optimise the process in real-time, ensuring quality standards and reducing the final financial loss.

In the olive oil extraction industry, the increasing demand for non-invasive measurement of oil content in by-products and waste remains a challenge for the process. Moreover, there is a need for studies performed directly at full-scale plants on real and uncontrolled samples to contribute to the ongoing discussion on real-scale applicability. The olive oil sector is in fact interested in new user-friendly systems for rapid analysis that can be performed directly on-line on the milling plant with the objective of using information from sensors to optimise the process. In the mill, the oil loss in by-products and waste is estimated to range between 10% and 15% of the total oil content in the olives. Such losses could occur for various reasons up to 20%–25% without the miller noticing, causing considerable economic losses. Automation in the olive milling industry is a key point for obtaining a high quality product through optimising process yields at low cost and thus a close cooperation between research teams and the industrial sector is necessary.

4. Conclusions

This study tested the applicability of vis/NIR spectroscopy directly at full-scale plants as a rapid technique for the prediction of the oil content at different points in the oil extraction process. A specific aim was to use spectral data sets covering different stages of the oil production processes in the calculation of models. Overall, the results are encouraging, except for measures on intact olives. Good results were obtained using combined datasets with paste and pomace samples, for both the vis/NIR and NIR spectral ranges.

The tested Device II was equipped with an immersion probe particularly suitable for liquid samples, and therefore good predictive models in validation (R^2 ranged 0.77–0.82) were calculated for the oil content prediction from the paté data set. Using the combined dataset with the paste and pomace samples, models with R^2_{val} in the range 0.89–0.92 were obtained, for both the measured spectral ranges, allowing us to envisage the possibility, in a future operative phase, of employing combined models usable in these two different stages of the process.

Regression models obtained could be the starting point to design a real-time prediction of crucial indices to support specific requirements of the process, considering that the tested devices are already designed for process applications and are built to be easily installed on process lines. The key aspect for the industrial application of these technologies is to get robust calibration models. Further studies will be needed because a full-scale application requires a lot of samples in order to develop models that can be used in a production context. This factor is decisive and partly explains why vis/NIR and NIR spectroscopy are still only installed to a limited extent in the industries in question. To help real scale calibration, the chemometric tools of devices must provide self-learning model calibration systems that involve the use of the acquired spectra directly on the process to strengthen models by expanding the data sets over time.

REFERENCES

Altieri, G., Di Renzo, G. C., & Genovese, F. (2013). Horizontal centrifuge with screw conveyor (decanter): Optimization of oil/water levels and differential speed during olive oil extraction. *Journal of Food Engineering, 119*(3), 561–572.

Armenta, S., Moros, J., Garrigues, S., & Guardia, M. D. L. (2010). The use of near-infrared spectrometry in the olive oil industry. *Critical Reviews in Food Science and Nutrition, 50*(6), 567–582.

Barros, A. S., Nunes, A., Martins, J., & Delgadillo, I. (2009). Determination of oil and water in olive and olive pomace by NIR and multivariate analysis. *Sensing and Instrumentation for Food Quality and Safety, 3*(3), 180–186.

Beghi, R., Giovenzana, V., Civelli, R., Cini, E., & Guidetti, R. (2013). Characterisation of olive fruit for the milling process by using visible/near infrared spectroscopy. *Journal of Agricultural Engineering, 44*, 56–61.

Bellincontro, A., Caruso, G., Mencarelli, F., & Gucci, R. (2013). Oil accumulation in intact olive fruits measured by near infrared spectroscopy—acousto-optically tunable filter. *Journal of the Science of Food and Agriculture, 93*(6), 1259–1265.

Bendini, A., Cerretani, L., Di Virgilio, F., Belloni, P., Lercker, G., & Toschi, T. (2007). In-process monitoring in industrial olive mill by means of FT-NIR. *European Journal of Lipid Science and Technology, 109*, 498–504.

Casale, M., Casolino, C., Ferrari, G., & Forina, M. (2008). Near infrared spectroscopy and class modelling techniques for the geographical authentication of Ligurian extra virgin olive oil. *Journal of Near Infrared Spectroscopy, 16*, 39–47.

Cherubini, C., Migliorini, M., Mugelli, M., Viti, P., Berti, A., Cini, E., et al. (2009). Towards a technological ripening index for olive oil fruits. *Journal of the Science of Food and Agriculture, 89*, 671–682.

Gallardo, L., Osorio, E., & Sanchez, J. (2005). Application of near infrared spectroscopy (NIRS) for the real-time determination of moisture and fat contents in olive pastes and wastes of oil extraction. *Alimentacion, Equipos y Tecnologia, 24*(206), 85–89.

García Sánchez, A., Ramos Martos, N., & Ballesteros, E. (2005). Comparative study of various analytical techniques NIR and NMR spectroscopies, and Soxhlet extraction for the determination of the fat and moisture content of olives and pomace obtained from Jaén Spain. *Grasas y Aceites, 56*, 220–227.

Giovenzana, V., Beghi, R., Civelli, R., Marai, S., & Guidetti, R. (2015). Postharvest characterization of olive oil fruits texture by NIR and vis/NIR spectroscopy. *Chemical Engineering Transactions, 44*, 61–66.

Huang, H., Yu, H., Xu, H., & Ying, Y. (2008). Near infrared spectroscopy for on/in-line monitoring of quality in foods and beverages: A review. *Journal of Food Engineering, 87*(3), 303–313.

Kavdir, I., Buyukcan, M. B., Lu, R., Kocabiyik, H., & Seker, M. (2009). Prediction of olive quality using FT-NIR spectroscopy in reflectance and transmittance modes. *Biosystems Engineering, 103*(3), 304–312.

Lachenmeier, D. W. (2007). Rapid quality control of spirit drinks and beer using multivariate data analysis of Fourier transform infrared spectra. *Food Chemistry, 101*(2), 825–832.

Leone, A., Romaniello, R., Zagaria, R., Sabella, E., De Bellis, L., & Tamborrino, A. (2015). Machining effects of different mechanical crushers on pit particle size and oil drop distribution in olive paste. *European Journal of Lipid Science and Technology, 117*(8), 1271–1279. https://doi.org/10.1002/ejlt.201400485.

Leone, A., Romaniello, R., Zagaria, R., & Tamborrino, A. (2015). Mathematical modelling of the performance parameters of a new decanter centrifuge generation. *Journal of Food Engineering, 166*, 10–20.

Marquez, A. J., Díaz, A. M., & Reguera, M. P. (2005). Using optical NIR sensor for on-line virgin olive oils characterization. *Sensors and Actuators B: Chemical, 107*(1), 64–68.

McGlone, V., Jordan, R., & Martinsen, P. (2002). Vis/NIR estimation at harvest of pre- and post-storage quality indices for 'RoyalGala' apple. *Postharvest Biology and Technology, 25*, 135–144.

Muik, B., Lendl, B., Molina-Díaz, A., & Pérez-Villarejo, L. (2004). Determination of oil and water content in olive pomace using near infrared and Raman spectrometry. A comparative study. *Analytical and Bioanalytical Chemistry, 379*(1), 35–41.

Nenadis, N., & Tsimidou, M. Z. (2017). Perspective of vibrational spectroscopy analytical methods in on-field/official control of olives and virgin olive oil. *European Journal of Lipid Science and Technology, 119*(1), 1600148.

Nicolaï, B., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K., et al. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology, 46*, 99–118.

Ortega, J. B., Gila, D. M. M., Puerto, D. A., García, J. G., & Ortega, J. G. (2016). Novel technologies for monitoring the in-line quality of virgin olive oil during manufacturing and storage. *Journal of the Science of Food and Agriculture, 96*, 4644–4662.

Porep, J. U., Kammerer, D. R., & Carle, R. (2015). On-line application of near infrared (NIR) spectroscopy in food production. *Trends in Food Science & Technology, 46*(2), 211–230.

Salguero-Chaparro, L., Baeten, V., Fernández-Pierna, J. A., & Peña-Rodríguez, F. (2013). Near infrared spectroscopy (NIRS) for on-line determination of quality parameters in intact olives. *Food Chemistry, 139*(1), 1121–1126.

Stella, E., Moscetti, R., Haff, R. P., Monarca, D., Cecchini, M., Contini, M., et al. (2015). Recent advances in the use of non-destructive near infrared spectroscopy for intact olive fruits. *Journal of Near Infrared Spectroscopy, 23*(4), 197–208.

Tamborrino, A., Leone, A., Romaniello, R., Catalano, P., & Bianchi, B. (2015). Comparative experiments to assess the performance of an innovative horizontal centrifuge working in a continuous olive oil plant. *Biosystems Engineering, 129*, 160–168.

Williams, P., & Norris, K. (2001). *Near infrared technology in the agriculture and food industries*. St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc.