agronomy

MDPI

Article

In-Field Assessment of Olive Fruit Quality Using a Low-Cost Multispectral Sensor and ANN Models

Miguel Noguera [1,*], Borja Millán [2], Arturo Aquino [1] and José Manuel Andújar [1]

[1] Centro de Investigación en Tecnología, Energía y Sostenibilidad (CITES), Universidad de Huelva, 21819 Huelva, Spain; arturo.aquino@diesia.uhu.es (A.A.); andujar@diesia.uhu.es (J.M.A.)
[2] Departamento de Ingeniería Eléctrica, Electrónica, Informática y de Sistemas, Universidad de Oviedo, 33203 Gijón, Spain; millanborja@uniovi.es
* Correspondence: miguel.noguera@dcaf.uhu.es

Abstract

Optimizing harvest time and oil production requires accurate olive fruit quality characterization. Traditional chemical methods are costly and tedious, leading to poor monitoring resolution and reliance on subjective visual assessments. While spectroscopy offers a non-destructive alternative, standard equipment remains complex and prohibitively expensive for smallholder farmers. To address this, we propose a methodology using a custom-made, low-cost multispectral device. Built upon the AS7265x board, the system acquires 18 spectral bands in the visible and near-infrared range (410–940 nm). We used these spectral data to feed artificial neural network (ANN) models for estimating the quality of intact olives. During a two-season field experiment, we monitored ripening to acquire spectral signatures and ground-truth values for oil content per fresh weight (OCFW), oil content per dry matter (OCDM), moisture (M), and titratable acidity (TA). External validation showed high accuracy for OCFW ($R^2p = 0.86$), OCDM ($R^2p = 0.86$), and M ($R^2p = 0.89$), proving the system's reliability. However, TA estimation showed lower performance ($R^2p = 0.21$), indicating limited spectral correlation. These findings pave the way for affordable, real-time smart farming tools for olive quality monitoring.

Keywords: olive fruit quality; multispectral spectroscopy; low-cost sensor; in-field assessment; artificial neural networks (ANN)

1. Introduction

Olive crop extension exceeds 12 million hectares around the world. Olive growing is linked to the Mediterranean climate conditions, as most cultivars require a specific temperature regime during the winter rest for a suitable flowering. This limits the optimal area for olive growing approximately between the latitudes of 30° and 45° [1]. Thus, the primary producers of olives are Spain, Italy, Greece, and Portugal, followed by other Mediterranean countries such as Tunisia, Morocco, Algeria, Egypt, and Turkey. However, over the last few decades, there have emerged producing countries from lower latitudes such as Argentina, Peru, Chile, and Australia. This competitive scenario has led to a transformation from the traditional olive orchards to hedgerow systems characterized by high (>1500 trees ha $^{-1}$), or super high plant densities (2116 trees ha $^{-1}$) [2]. These new crop settings involve modern resources like drip irrigation systems, no-tillage techniques, and quick-growing young olive plantings in rows, which enable labor mechanization (harvest, prune, fertilize, phytosanitary application, etc.). This new paradigm results in an

Academic Editors: Claudia Savarese, Angelo Cardellicchio and Vito Renò

Received: 12 May 2026
Revised: 16 June 2026
Accepted: 17 June 2026
Published: 19 June 2026

increase in yield per hectare while reducing management costs, thus making the crop more profitable. This modernization of olive growing has incentivized investigations aimed at exploring the applicability of precision agriculture techniques, such as the development of sensors capable of assessing parameters of interest for olive management. In this sense, the investigation of new methodologies for olive fruit quality assessment has attracted particular interest in recent decades [3,4].

The decision regarding the optimal harvest time is critical for olive producers, as it must balance both oil yield and quality. The main indicator that olives have reached this stage is the completion of oil biosynthesis [5], and the most widely accepted tool in the olive sector for identifying this stage is the maturity index, which is based on the colour of the fruit's skin and flesh [6]. However, several agronomic factors, such as climate conditions, amount of fruit on the tree, nitrogen availability, and cultivar differences, can weaken the correlation between fruit colour and chemical composition [7,8]. These limitations make the assessment of olive ripeness based solely on skin and flesh colour sometimes affected by subjectivity [9].

Oil content, along with moisture and titratable acidity, is a key parameter used by mills to classify olives into different commercial quality categories. On the other hand, the evolution of these parameters during olive ripening can occur at different rates across a crop, leading to a heterogeneous quality status. This variability can be managed through targeted agronomic practices [8]. Therefore, in-field monitoring of these parameters becomes especially relevant for olive growers. Traditionally, the determination of these parameters is conducted by standard analytical methods. A usual methodology for acidity determination in olive fruits relies on the titration of olive oil [10]. Oil content can be determined by Soxhlet or other techniques such as nuclear magnetic resonance (NMR) [11]. Moisture is normally determined by using a drying oven and Karl Fischer titration [12]. These classical chemical methods have a negative environmental impact due to the use of polluting reagents and solvents. Additionally, they require expert personnel and specialized laboratory equipment, leading to high costs. Furthermore, these methods are destructive, labor-intensive, and time-consuming, often requiring several days to deliver results. This delay limits their application in real-time decision-making and reduces the number of sample points that growers can feasibly consider, resulting in poor spatial resolution. Similarly, these constraints restrict the frequency of sampling campaigns throughout the season, leading to low temporal resolution. Consequently, growers often rely on homogeneous harvesting strategies based on subjective criteria such as intuition and visual interpretation. In contrast, real-time field measurements of these chemical parameters could provide valuable insights for optimizing harvest timing and overall orchard management. First, they would enable site-specific treatments during growth and early ripening phases to mitigate quality heterogeneity. Additionally, growers could adopt targeted harvest strategies to diversify production, enhancing their economic returns and increasing the capacity of the agri-food sector to develop higher-value products.

In response to the described context, the present work proposes a non-destructive methodology implementing a handheld VIS-NIR multispectral device for estimating quality indicators of olive fruits under field conditions. The agronomical applications of spectroscopy rely on the interaction between electromagnetic radiation and vegetal tissues. Each vegetal tissue (leaves, fruits, stems, etc.) has a different chemical composition, which also varies along the vegetative cycle. When an electromagnetic beam interacts with a vegetal tissue, it is modified in a specific way, which is determined by its chemical composition. This modification can be assessed by analyzing the three processes of absorption, reflection, and transmission. So, as no two organic compounds have the same characteristics, a compound can be identified by characterizing its absorption spectrum and matching it with a

database. Spectroscopy-based techniques have been widely investigated in the precision farming sector [13]. However, the research focused on olive crops and concretely in the quality status assessment of olive fruits is relatively scarce.

The first approaches exploiting spectroscopy for olive quality estimation focused on the use of NIR spectrometers applied to olive by-products such as olive paste and pomace [14–16]. These techniques demonstrated accuracy comparable to that of standard chemical methods, with the added advantages of requiring less intensive sample handling, reduced use of reagents and solvents, and faster processing times compared to official procedures. However, they still involved sample destruction. Subsequently, numerous studies have highlighted the potential of VIS-NIR spectroscopy for assessing various quality parameters of intact olive fruits (maturity index, firmness, soluble solid content, phenolic compounds, dry matter, moisture, acidity and fat content) under controlled laboratory conditions [3,4,17–22]. An additional challenge associated with these methodologies is the acquisition of spectral data under field conditions, as measurements can be affected by variable lighting conditions inherent to real-world scenarios. Nevertheless, this also offers a technical advantage for practical applications, as it supposes immediate access to information, whose usefulness justifies the pursuit of an effective analytical methodology. The feasibility of assessing quality parameters of olive fruits based on spectral data acquired in field conditions has been previously reported. In this regard, Garcia & León [23] proposed a methodology based on a portable AOFT-NIR spectrophotometer combined with chemometrics methods (partial least squares regression (PLSR)) for the determination of oil and moisture contents in intact olive fruits. They compared the performance of PLSR models fed with spectral data acquired under field and laboratory conditions, concluding that although slightly better results were obtained under laboratory conditions, the results obtained in the field were also accurate enough. More recently, Fernández-Espinosa et al. [3] compared the spectra obtained with a NIR spectrometer from olive samples at different stages: on the tree before harvesting, immediately after harvesting on the ground under the tree, and in the laboratory at 4, 24, and 48 h, as well as one week after harvesting. The most significant differences in the spectra were observed beyond 24 h after harvesting. The study concluded that the olive spectrum can be reliably captured both in the field and in the laboratory, provided that the measurement is not delayed beyond 24 h after harvesting. Despite the proven potential of spectroscopy, the lack of low-cost, user-friendly, and field-robust solutions has prevented widespread adoption, particularly among smallholder farmers.

In recent times, the microelectronic industry has made significant advancements, offering components with enhanced features at reduced costs. These upgrades have aroused an increasing interest in the use of low-cost spectral sensors for diverse agronomical applications [24–30]. Taking advantage of this context, the present research proposes a methodology based on a low-cost, VIS-NIR multispectral device (18 bands between 410 and 940 nm) and artificial neural network (ANN) models to estimate quality parameters of intact olive fruits under field conditions. The hardware of the proposed handheld device is based on a VIS-NIR multispectral sensor (AS7265x, AMS AG, Graz, Austria) commercially available. In a previous work, based on an initial prototype adapted to operate under laboratory conditions, the same authors demonstrated the suitability of this component to assess quality parameters of olive fruits [31]. Afterwards, this sensor was implemented on a second prototype with a specific hardware intended to operate under field conditions, which was evaluated for grape quality evaluation with promising results [32]. Given these backgrounds, the goal of this research was to evaluate the potential of the developed handheld VIS-NIR multispectral device for assessing quality parameters of intact olive fruits under field conditions. To this end, the device was tested in a field experiment conducted over two consecutive seasons, during which it was used to characterize the reflectance

spectrum of olive samples collected periodically throughout the ripening process. These samples were also analysed using standard chemical methods to determine quality indicators. The spectral information acquired by the sensor served as input variables for training ANN models, using the chemical analysis results as a reference. This study compiles the results obtained from that experiment.

2. Materials and Methods

2.1. Device Description

The aim of this research was to develop a reliable, low-cost, and user-friendly tool capable of quantitatively assessing the quality status of olive fruits under field conditions. The resulting device (Figure 1) integrates several commercially available components connected via custom-designed printed circuit boards (PCBs) and housed in a custom 3D-printed enclosure.

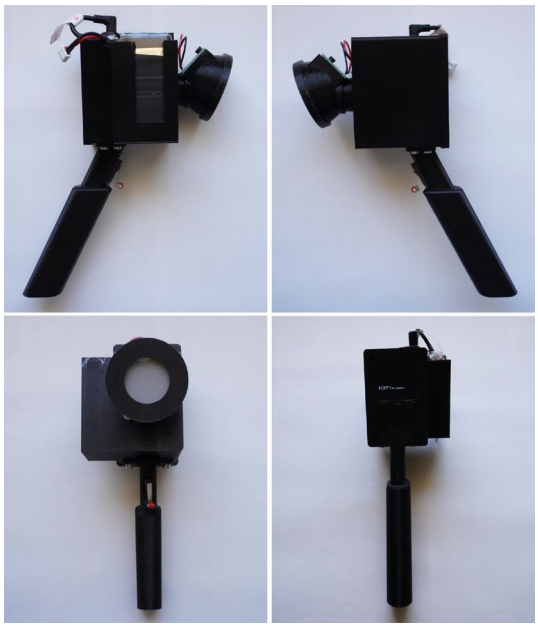


Figure 1. Developed a prototype from four different views. The three parts of the framework (handle, dome, and box-type enclosure), the trigger button, the OLED screen, and the external battery are shown.

The core of the device is an AS7265x development board (AMS AG, Graz, Austria). This board includes three main chips: AS72651, AS72652, and AS72653, each equipped with six optical filters, making them sensitive to six distinct spectral bands. Combined, they form an 18-channel VIS-NIR multispectral sensor with a spectral response range from 410 nm to 940 nm and a full width at half maximum (FWHM) of 20 nm.

As an illumination source, the device uses an array of three IR broadband emitters (OSLON P1616 SFH 4737, OSRAM, Munich, Germany). Having an integrated light source enables the device to characterize the incident radiation at the time of measurement, thereby ensuring the comparability of spectral signatures.

Additionally, the device features a 1.3-inch OLED display (128 × 64 pixels) (SparkFun Electronics, Niwot, CO, USA), which provides the user with key information during operation, including configuration settings, file names, the number of measurements taken, and the current system status.

All components communicate through an Arduino MKRZero Board (Arduino LLC, Monza, Italy), selected for its compact size, low power consumption, affordability, and built-in SD card slot.

The system software was developed using the Arduino IDE (version 1.8, Arduino LLC) and designed with a focus on user-friendliness. Upon powering on, the device automatically generates a new data file to store the upcoming measurements. It then enters an idle state, awaiting user input via the trigger button.

When the trigger is pressed, the Arduino board sends a command to the sensor and activates the LEDs, whose intensity can be configured through software. The device then captures the spectral data, which, along with relevant metadata such as file names, is saved to an SD card for later analysis (Figure 2). After the acquisition, the device returns to its idle state, ready for the next measurement.

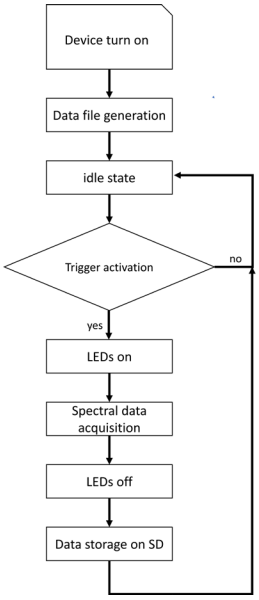


Figure 2. Flow-chart diagram of the device operation.

The controller board can also be connected to a PC to configure internal parameters of the device, such as exposure time, gain, lighting duration, and LED current.

The device is powered by a 2S LiPo (Lithium-ion Polymer) battery connected directly to the controller board. All components are housed within a custom-designed enclosure, created using FreeCAD (version 0.16, The FreeCAD Team) and manufactured via 3D printing with biodegradable polylactic acid (PLA) filament.

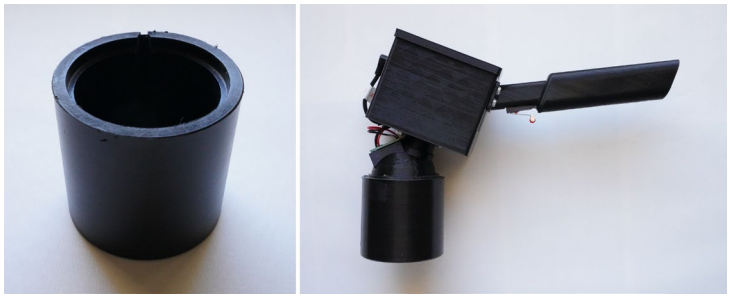
The enclosure consists of three main parts:

- A handle that houses the trigger button.
- A dome that supports the light source and integrates a light diffusion film (OptSaver L-9960, Kimoto LDT, Zurich, Switzerland) placed in front of the sensor to homogenize the radiation reflected by the sample. This dome was designed to position the sample at a 45° angle relative to both the light source and the sensor, optimizing measurement accuracy.
- A box-type structure that contains the AMS AS7265x development board, the Arduino MKR Zero, the interconnection board, and a lid that seals the system for in-field operation. The OLED display is mounted on this lid.

All elements of the device are contained within this custom enclosure, except for the battery. The battery was intentionally installed externally to prevent thermal interference with the sensor, ensuring measurement accuracy, and to facilitate quick and easy replacement during fieldwork.

A more exhaustive description of the hardware and software of the device can be found in Noguera et al. 2022 [32]. In the present research, the device was adapted for measuring

the spectral signature of intact olive fruits in field conditions. In this regard, a custom container was designed to fit tightly with the dome of the device. This container enables the secure placement of olive samples during measurement, effectively isolating them from ambient radiation interference. The container has a suitable volume to accommodate approximately 50 g of olive fruits. An image of the developed device, along with the customized sample container, is shown below (Figure 3).



(**a**) (**b**)

Figure 3. Images of the measuring container alone (**a**) and fitted with the device's dome during measurement (**b**).

Evaluation of Environmental Radiation Interference

In order to evaluate the capabilities of the proposed system to avoid the interference of environmental radiation, a specific experiment was carried out. This experiment consisted of exposing the device to different levels of environmental radiation. Thus, three reference scenarios were defined. Firstly, the device was used in a dark room, completely isolated from environmental radiation. Additionally, the device was used in two differentiated in-field environmental illumination conditions: sheltered by the shade of an olive tree, and completely exposed to the sun's radiation. To isolate the environmental radiation as the unique variable to be considered in the comparison, the measuring container remained empty in this experiment. Ten measurements were acquired in each scenario, which were calibrated using the reflectance of a calibrated surface, as it is described in Section 2.3.1.

To determine the statistical significance of the levels of environmental radiation on the performance of the proposed device, an analysis of variance (ANOVA) of a single factor with a significance level $\alpha = 0.05$ was made on every single band. The ANOVA analysis has the ability to estimate the effect of any treatment by taking the difference between the mean of the observations that receive the treatment and the general mean.

2.2. Calibration and Testing of the Device's Capability for Assessing Olive Fruit Quality Parameters

A field experiment was carried out to evaluate the capability of the proposed device for quantitatively assessing quality indicators of intact olive fruits.

2.2.1. Study Site Description

The study site was a commercial olive orchard (*Olea europaea* L., cv. Picual), managed by Nuestra Señora de la Oliva, S.C.A., and placed in the province of Huelva, Andalusia, Spain (37°20′28.96″ N and 7°01′54.98″ O). It has a plant spacing pattern of 7 × 7 m, which corresponds to a traditional olive orchard. This farm covers an area of more than 200 hectares; however, for this experiment, a one-hectare plot was delimited (highlighted in red in Figure 4). This delimitation was intended to ensure a homogeneous ripening pace within the experimental plot, thereby isolating the natural evolution of quality parameters during ripening as the sole source of variability.

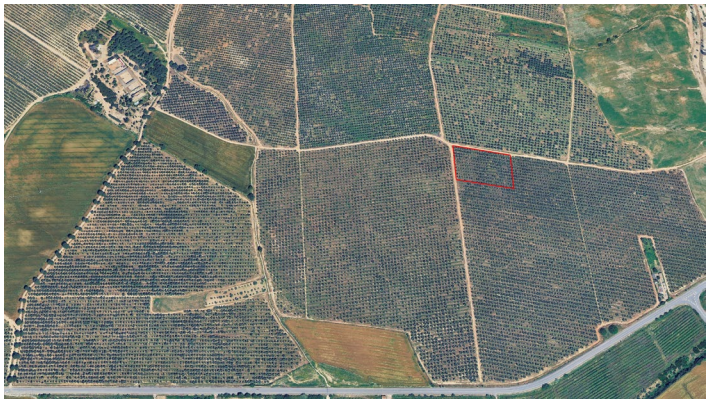


Figure 4. Bird eye image (Google Earth Pro satellite view) of the cultivar used for experimentation; the red square represents the delimited area.

2.2.2. Experimental Field Protocol Design

During two consecutive seasons, the experimental field was periodically sampled to cover the entire ripening process. The first sampling was carried out after the post-summer fruit growth stage, just before the onset of veraison (October) and was repeated weekly until harvest (December). Each sample consisted of approximately 200 g of olives collected from the same branch of a randomly selected tree. To minimize intra-tree variability caused by differing canopy microclimates (such as varying sunlight exposure), sampling was consistently performed on branches located at the outer canopy, at a standard operator height (approximately 1.5 m). Collecting the entire sample batch from a single branch ensures that the fruits within each 200 g unit share a highly homogeneous physiological and ripening status, thereby mitigating the risk of large intra-sample variations. The spectral signature of each sample was acquired on the tree immediately after harvesting, after which the samples were packaged, labeled, and refrigerated for transport to the laboratory for reference analysis.

The field experiments began in the 2023 season. However, due to unpredictable climatic conditions that caused a sharp maturation peak between the second and third sampling sessions, a gap in the distribution of the assessed quality parameters was identified (see Section 3.2). In machine learning applications for agriculture, models trained on data from a single season are often prone to overfitting and may struggle to generalize when exposed to natural inter-annual climatic variations. Therefore, to ensure a rigorous experimental design and build a highly robust ANN training dataset, the sampling protocol was actively repeated during the 2024 season. This strategy successfully filled the missing range of quality parameters while incorporating valuable inter-seasonal variability. As a result, the final dataset comprised 86 samples collected in 2023 and 60 in 2024, totalling 146 samples, providing a comprehensive and generalized foundation for model training.

2.2.3. Spectral Data Collection

The spectral signature of the samples was acquired under field conditions, immediately after being picked. As soon as the olive fruits were collected, they were put in the specific container (Figure 5a), the container was fitted with the device dome (Figure 5b), and three consecutive spectral captures were taken; the three spectral readings were subsequently averaged. The fruits were then packaged, labelled, and chemically analyzed as described in Section 2.2.4. to generate the reference information. The chemical methods required a minimum amount of 200 g of olive fruits, so four containers of fruit (~50 g each) were required to configure a sample unit (~200 g). Thus, 12 spectral acquisitions were performed per sample unit (3 captures/container $\times$ 4 containers). This bulk-averaging strategy, combined with the light diffusion film and the fixed measuring geometry of the

device dome, was specifically designed to overcome the morphological and structural heterogeneity of individual olive fruits (such as irregular shapes and variations in skin thickness or pulp structure). By integrating the spectral response of approximately 200 g of fruit across multiple captures, individual physical differences are effectively neutralized, ensuring the consistency, comparability, and representativeness of the final spectral signature used to feed the models.

(**a**) (**b**)

Figure 5. Images of the measuring container loaded with olive fruits (**a**) and fitted with the device's dome during measurement (**b**).

Once every five samples, three captures of a characterized reflectance surface were taken (Labsphere, Inc, North Sutton, NH, USA) (53% calibrated reflectance). The average signal of the three captures was used as a reference to normalize the spectral signatures of the following five samples registered.

2.2.4. Reference Analysis

In this research, the oil content (OC), moisture (M), and titratable acidity (TA) were used as target parameters, as they are widely used for classifying olive fruits into different commercial qualities. The reference analysis of these parameters was made in the laboratory, in accordance with official methodologies. The determination of moisture was carried out through the drying method in an oven at 105 °C (ISO662:2016) [12]. According to this method, a 25 g portion of the sample is placed in a porcelain capsule and dried in the oven at 105 °C for 6 h. Then, the sample is cooled in a desiccator, weighed, and returned to the oven, repeating these operations until, between two consecutive weightings, the variation in weight loss of moisture is less than 0.02 g. At this point, the difference between the initial and final weight is considered the moisture content of the sample. Oil content was determined in its two variants: oil content per dry matter (OCDM) and oil content per fresh weight (OCFW). Reference analysis of OCFW was performed using the Soxhlet methodology (UNE 55030:1961) [11]. This method involves introducing the dried sample, used for moisture determination, into a Soxhlet extractor, where the fat is extracted by exposing the sample to n-hexane for 4 h. Then, the sample is placed into the oven again at 105 °C to remove the traces of solvents. The amount of oil recovered is used to determine the OCFW and, from this measure, the OCDM is calculated according to the following Equation (1):

$$\text{OCDM} = \frac{\text{OCFW}}{100 - \text{M}} \text{ x } 100 \quad (1)$$

The TA of the olive samples was determined by means of titration, according to the international standard methodology (UNE-EN ISO 660:2020) [10]. This method consists of a titration of 1–3 g of oil with ethanolic potash in a 5 mL microburette.

2.3. Methodology for Olive Quality Indicators Estimation from Multispectral Data

2.3.1. Data Pre-Processing

During the field experimentation, 12 spectral readings were registered per olive sample (3 captures/container × 4 containers). Thus, the first step of the data pre-processing was to average those 12 spectral signatures from each olive sample. As a result, each sample was summarized to 18 reflectance data. Then, these data were normalized using as a reference the spectral signature of a calibrated reflectance surface (53%) (Labsphere, Inc, North Sutton, NH, USA), which was captured once every five samples. The 18 reflectance signals of the calibration surface were used as references for the spectral signature of the subsequent five samples according to the following Equation (2):

$$Rcal_{wl} = \frac{R_{wl}}{Rref_{wl}/53} \quad (2)$$

where R_{wl} is the reflectance value measured for a given spectral band in a capture of a sample, $Rref_{wl}$ is the reflectance value measured for that spectral band in the previous capture of the calibrated reflectance surface, and $Rcal_{wl}$ is the corrected value of reflectance in the sample of the given band. After this normalization, the data were transformed to fall in the interval [0, 100], thus making the reflectance values of the 18 spectral bands mathematically comparable. Additionally, the pre-processed spectra were smoothed for noise removal by applying the Savitzky–Golay method [33]. After considering different parameter settings, the width of the selection window was adjusted to five points and a second-order polynomial was used to fit the data. No derivatives of the data were made, as only a simple smoothing was intended.

2.3.2. Development of Prediction Models

The pre-processed reflectance values of the 18 spectral bands acquired by the sensor were used as input variables to train four Artificial Neural Networks (ANN) aiming to estimate OCFW, OCDM, M, and TA, respectively. The software used for data processing and ANN training was MATLAB (version R2024b, The MathWorks Inc., Natick, MA, USA). The selected ANN architecture consisted of a two-layer feed-forward perceptron with back-propagation. Specifically, the network was composed of a hidden layer with ten neurons, an input layer with eighteen neurons (corresponding to the corrected reflectance values), and an output layer with one neuron (the respective target parameter). The activation function for the hidden layer was sigmoid, while a linear transfer function was used for the output layer. The Levenberg–Marquardt backpropagation algorithm was employed for network training.

Model validation was performed using a random sampling (RS) strategy. To this end, the complete dataset was randomly partitioned into three subsets: 70% for training (n = 102), 15% for internal validation (n = 22), and 15% for external validation or testing (n = 22). The training set is presented to the network to adjust the weights based on the error minimization. The internal validation set is used to monitor the network's generalization capability and to halt training when performance stops improving (early stopping). Finally, the external validation set, which has no effect on the training process, provides an independent measure of the network's performance.

2.3.3. Criteria for Model Performance Evaluation

The performance of the estimation models was evaluated during both calibration and external validation stages using the following metrics: the coefficient of determination for calibration (R^2c) and external validation (R^2p), the root mean square error of calibration (RMSEC) and prediction (RMSEP), the standard error of calibration (SEC) and prediction (SEP), and the ratio of performance to deviation (RPD).

The coefficient of determination (R^2) was used to assess the fit between the actual and predicted target parameters. Specifically, R^2 indicates the percentage of the variance in the target variable that is accounted for by the model's response, and is calculated as follows:

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \overline{y})^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2} \quad (3)$$

where y_i is the reference value, $\hat{y}_i$ is the predicted value, $\overline{y}$ is the mean of the reference values, and n is the number of samples. An R^2 value between 0.50 and 0.65 suggests that more than 50% of the variance is explained, enabling discrimination between high and low levels. Values between 0.66 and 0.81 indicate approximate quantitative predictions, whereas values higher than 0.82 reveal good predictive performance, improving as R^2 approaches 1 [34].

The RMSE served as the primary metric for evaluating model accuracy, quantifying the overall magnitude of the prediction errors.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \quad (4)$$

By aggregating the squared differences between the predicted olive quality parameters and the reference values, the RMSE incorporates both systematic bias and random variation, thereby penalizing larger deviations. In contrast, the SE was calculated to assess model precision by measuring the dispersion of residuals around the regression line:

$$SE = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\hat{y}_i - y_i - bias)^2} \quad (5)$$

where $bias = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)$. Defined as the standard deviation of the prediction errors, the SE isolates the random variability of the ANN model, independent of systematic bias. Consequently, a lower SE indicates that the predicted values are clustered more closely to the linear fit.

Furthermore, the coefficient of variation in the RMSE and SE of calibration and prediction (rRMSEC, rRMSEP, rSEC and rSEP), which are the result of normalizing the original parameters by the mean value of the reference data ($\overline{y}$) in the respective dataset, were also considered:

$$rRMSE(\%) = \frac{RMSE}{\overline{y}} \times 100 \quad (6)$$

$$rSE(\%) = \frac{SE}{\overline{y}} \times 100 \quad (7)$$

This allows for the avoidance of ambiguity, facilitating the comparison between datasets or models with different scales. In this sense, RPD is the ratio between the standard deviation (SD) of the external validation set and the SEP, so it is also dimensionless.

$$RPD = \frac{SD}{SEP} \quad (8)$$

RPD indicates the accuracy of the prediction compared to the average composition of all samples. For this metric, it is usually accepted that models with an RPD smaller than 1.5 are not suitable, while those showing RPD values between 1.5 and 2.0 can discriminate low from high values of the response variable, while RPDs greater than 2.0 exhibit a very good predictive performance, which can be considered excellent when RPDs exceed 3.0 [35,36].

3. Results

3.1. Effect of Environmental Radiation on Device Performance

Figure 6 shows the average spectral reflectance for each band captured by the sensor, obtained from 10 measurements of a constant target under three different measurement conditions with varying levels of ambient radiation: a dark room, an open field fully exposed to ambient radiation, and a partially shaded area under trees. The overlap of the mean spectral signatures suggests that the proposed measuring system effectively isolates measurements from environmental radiation interference. This visual assumption was corroborated by the results of an analysis of variance (ANOVA) of a single factor, which found no significant differences ($\alpha = 0.05$) between the mean values of the measurements taken under each individual scenario and the overall mean across all three scenarios for all considered bands.

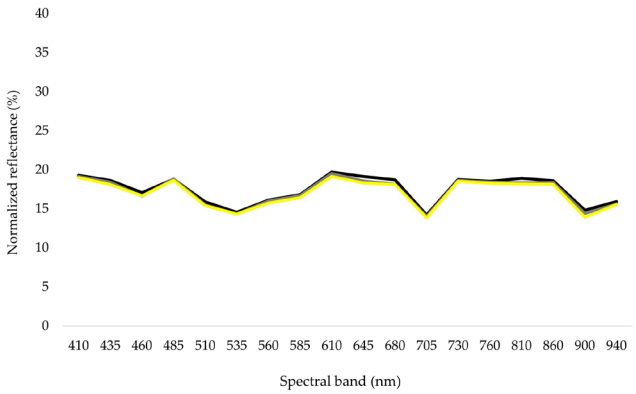


Figure 6. Mean reflectance response of the empty measuring container in a dark room (black), under field conditions under shadow (grey), and exposed to sun radiation (yellow).

3.2. Reference Values of Target Quality Parameters

Figure 7 represents the reference values of the four studied parameters (oil content per fresh weight (OCFW), oil content per dry matter (OCDM), moisture (M), and titratable acidity (TA)) for each individual sample, collected weekly over the two consecutive seasons ($n = 146$). Samples numbered 0 to 86 correspond to the first season, while the remaining samples were acquired during the following season.

Figure 7a,b illustrate a sharp increase in OCFW and OCDM values between the second and third sampling events of the first season. This resulted in a gap in the data distribution, which hindered the proper training of the ANN models. However, this issue was addressed in the second season, when additional samples were collected to cover the missing range. The discrete weekly sampling protocol, combined with this rapid ripening event and the integration of data from two distinct seasons, naturally results in the visual clustering of the data points observed in Figure 7. As a result, when considering both seasons together, OCFW ranged from 13.36% to 23.87%, with an average of 19.21% $\pm$ 2.68%, while OCDM varied between 32.70% and 51.73%, with an average of 44.55% $\pm$ 5.20% (Table 1).

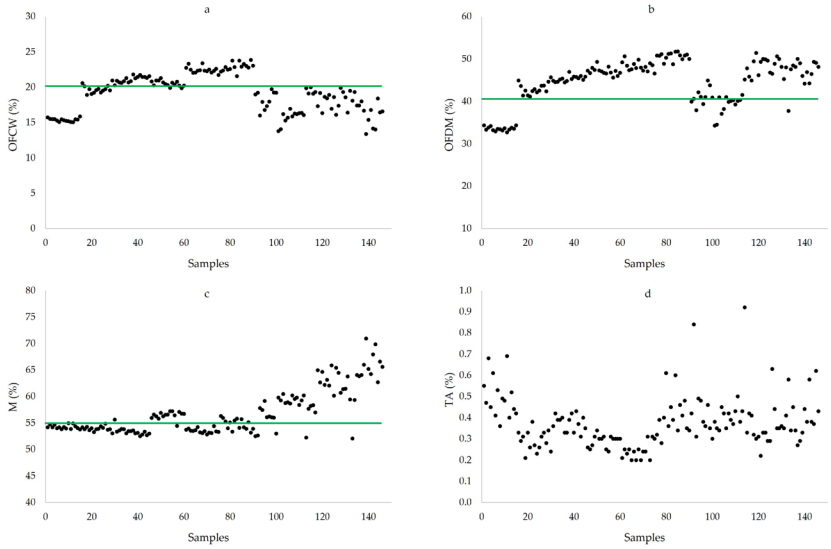


Figure 7. Evolution of the target quality parameters across all the collected olive samples (n = 146): (**a**) OCFW; (**b**) OCDM; (**c**) M; and (**d**) TA. The green line represents the recommended minimum levels of OCFW, OCDM, and M for harvest.

Table 1. Summary descriptive statistics of the olive fruit dataset related to OCFW, OCDM, M, and TA.

	Min	Max	Mean	SD
OCFW (%)	13.36	23.87	19.21	2.68
OCDM (%)	32.70	51.73	44.55	5.20
M (%)	52.04	70.94	56.79	4.17
TA (%)	0.20	0.82	0.37	0.11

Regarding moisture (Figure 7c), values remained relatively stable during the first season, with no significant variations across weekly samplings. This could be attributed to a period of drought stress, which was particularly severe that year. In contrast, during the second season, M values exhibited a steady increase over time. Overall, combining data from both seasons, M values ranged from 52.04% to 70.94%, with an average of 56.79% $\pm$ 4.17% (Table 1). The replication of the experimental protocol ensured a broad distribution of OCFW, OCDM, and M values, both above and below the minimum thresholds recommended for harvesting [8].

With respect to TA, values ranged from 0.20% to 0.82%, with an average of 0.37% $\pm$ 0.11% (Table 1). Figure 7d shows that TA exhibited significant variability, even among samples collected during the same sampling event, without a clear trend related to the ripening process. This could be explained by the fact that TA levels are highly influenced by fruit damage caused by various factors, such as hail, frost, insect bites, and phytopathogenic infections [8]. The acidity of the olive oil produced after the extraction process is a key parameter in determining its quality category. According to the virgin olive oil standard [37], oils with an acidity below 0.8% are classified as extra virgin olive oil, while virgin olive oil must have an acidity of less than 2%. In addition, oils with acidity levels above 2% are considered lampante oils, which must undergo refining processes to make them suitable for consumption [38]. The extraction process increases oil acidity compared to that of the fruit. Consequently, acidity levels measured in the fruit should be considered an indicator of fruit quality rather than a precise reference for the acidity level of the oil that will be produced. In this study, all samples were picked directly from trees. Therefore, the observed TA range was consistent with expectations.

3.3. Spectral Signature of Olive Samples

Prior to the development of the estimation models, a visual inspection of the spectral response of the olive samples was conducted. Visual features and patterns of the spectra conform with those previously reported in the literature for intact olive drupes [39]. Figure 8 represents the mean spectral signature of the samples corresponding to the 10th and above the 90th percentile of the studied parameters. Indeed, for the cases of OCFW (Figure 8a) and OCDM (Figure 8b), the overall shapes were rather similar at both ends of the histogram. Generally, a decrease in reflectance can be observed in the case of those samples with higher levels of OCFW and OCDM, compared to those samples with lower levels. Notwithstanding, the decrease in reflectance was higher for those samples with higher OCDM values. Especially noteworthy were the reflectance peaks between 585 and 900 nm in the case of samples with lower levels of OCFW and OCDM, compared to those with higher levels. In relation to M, the trend was similar, showing a decrease in reflectance linked to an increase in M levels. In this case, the decrease in reflectance extended to 435–560 nm. However, the reflectance decreases observed in the samples with high levels of M at 610 nm and 645 nm were less than that shown by the samples with elevated levels of OCFW and OCDM. Finally, there were no differences in the spectral signature of the samples related to TA level variations.

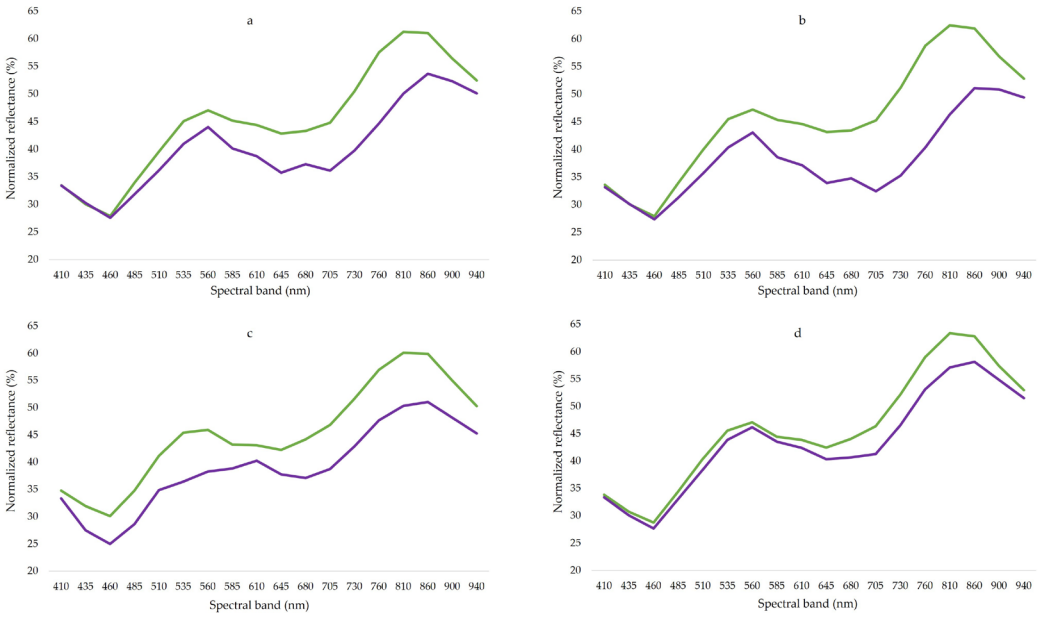


Figure 8. Mean reflectance response of olive samples corresponding to the 10th (green line) and above the 90th (purple line) percentile for each target parameter: (**a**) OCFW; (**b**) OCDM; (**c**) M; and (**d**) TA.

3.4. Performance of Estimation Models

The performance of the four ANN models was evaluated by means of a random sampling test (RS). The statistics obtained during the model calibration phase are detailed in Table 2, whereas Table 3 presents the statistics comparing the ANNs' output and the reference data in the external validation set.

Regarding the goodness of fit, the ANN models showed strong relationships with the reference values for OCFW, OCDM, and M, achieving coefficients of determination of prediction (R^2p) of 0.86, 0.86, and 0.89, respectively (Figure 9). In the calibration set, the coefficients (R^2c) were slightly lower in the case of OCFW ($R^2c = 0.84$) and OCDM ($R^2c = 0.85$), and slightly higher in the case of M ($R^2c = 0.90$). Conversely, the TA model exhibited a poorer performance, yielding an R^2p of 0.21 in external validation and an R^2c of 0.31 during calibration.

Table 2. R^2c, RMSEC, rRMSEC (%), SEC and rSEC (%) between reference values of OCFW, OCDM, M, and TA, measured by chemical methods, and those estimated based on ANN approaches in the calibration sets.

Calibration					
Parameters	**R^2c**	**RMSEC**	**rRMSEC**	**SEC**	**rSEC**
OCFW (%)	0.84	1.11	5.83	1.20	6.34
OCDM (%)	0.85	1.88	4.17	2.04	4.53
M (%)	0.90	1.45	2.55	1.57	2.77
TA (%)	0.31	0.10	26.28	0.11	28.56

Table 3. R^2p, RMSEP, rRMSEP (%), SEP, rSEP (%), and RPD between reference values of OCFW, OCDM, M, and TA, measured by chemical methods, and those estimated based on ANN approaches in the external validation sets.

External Validation						
Parameters	**R^2p**	**RMSEP**	**rRMSEP**	**SEP**	**rSEP**	**RPD**
OCFW (%)	0.86	1.29	6.78	1.30	6.85	2.38
OCDM (%)	0.86	2.74	6.25	2.52	5.75	2.47
M (%)	0.89	1.66	2.88	1.65	2.86	2.58
TA (%)	0.21	0.17	42.38	0.17	40.21	1.12

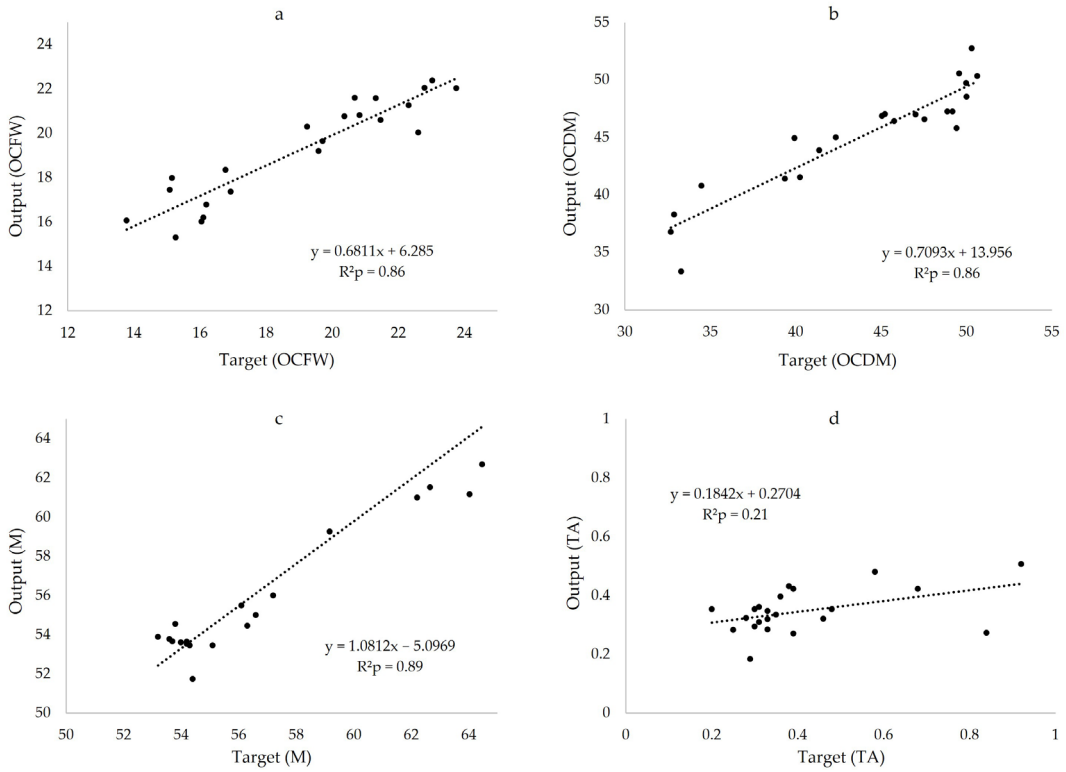


Figure 9. Linear regressions between the reference values measured by chemical methods and those estimated based on ANN approaches in the external validation set for: (**a**) OCFW; (**b**) OCDM; (**c**) M; and (**d**) TA.

In terms of accuracy, the computed root mean square errors revealed the magnitude of the total deviation. For the calibration set, RMSEC values were 1.11% for OCFW, 1.88% for OCDM, 1.45% for M, and 0.10% for TA (Table 2). In external validation, RMSEP values increased to 1.29% for OCFW, 2.74% for OCDM, 1.66% for M, and 0.17% for TA. To account for the different scales of the parameters, the relative errors were analysed. The rRMSEC

confirmed the high accuracy of the M model (2.55%), followed by OCDM (4.17%) and OCFW (5.83%), whereas the TA model showed a significantly higher relative error (26.28%). A similar hierarchy was observed in the validation set, where rRMSEP values were 2.88% (M), 6.25% (OCDM), 6.78% (OCFW), and 42.38% (TA).

With respect to precision, the SEC and SEP values followed closely the trends observed for RMSE. SEC values ranged from 0.11% (TA) to 2.04% (OCDM), while SEP values in the external validation set were 1.30% for OCFW, 2.52% for OCDM, 1.65% for M, and 0.17% for TA. When normalized, the relative standard error of prediction (rSEP) further highlighted the differences in model performance: the M model achieved the best precision (rSEP of 2.86%), followed by OCDM (5.75%) and OCFW (6.85%). In contrast, the TA model exhibited a high dispersion relative to its mean, with an rSEP of 40.21%.

Finally, the RPD values corroborated the classification of the models. The M model achieved the highest performance with an RPD of 2.58, followed closely by OCDM (2.47) and OCFW (2.38), indicating that these models are suitable for approximate quantitative predictions and screening. On the other hand, the TA model yielded an RPD of 1.12, which falls below the threshold of 1.5, confirming its unsuitability for reliable prediction purposes.

4. Discussion

Prior to the in vivo experiment, a preliminary test was conducted to evaluate the capability of the proposed device to avoid the influence of environmental radiation. To this end, the device was used to acquire the spectral signature of a constant target under three different scenarios with increasing levels of ambient radiation. The mean reflectance values for each of the 18 spectral bands captured by the sensor were compared across the three scenarios using a one-way ANOVA. The analysis revealed no statistically significant differences ($\alpha = 0.05$) in any of the spectral bands. These findings suggest that the performance of the proposed device is not affected by varying levels of ambient radiation, thereby supporting its reliability for in-field operation.

The field sampling began in the 2023 season. However, a sharp ripening progression between two sampling events resulted in a gap in the distribution of quality parameters. Since the range and distribution of the training dataset are critical factors influencing the performance of machine learning models, such as ANNs, a second round of field experiments was conducted during the 2024 season to address this gap. As a result, the combined dataset from both seasons exhibited a wide and homogeneous distribution of the target parameters, including values both above and below the recommended harvest thresholds [8]. Additionally, the inclusion of data from two consecutive growing seasons enhances the robustness and generalization capability of the estimation models. By incorporating inter-annual variability into the training dataset, the models avoid the common pitfall of overfitting to a single year's specific weather conditions, thereby ensuring a more realistic predictive performance.

The spectral range acquired by the proposed device (410–940 nm) encompasses both the VIS spectrum (410–760 nm) and a portion of the NIR region (810–940 nm). The observed decreases in reflectance associated with high levels of OCFW, OCDM, and M within the VIS domain may be attributed to increased absorbance resulting from variations in pigment content, including chlorophylls (a and b), carotenoids, and anthocyanins. During the ripening process, anthocyanins accumulate in the olive fruit as oil content increases. Concurrently, photosynthetic activity declines, leading to a progressive reduction in chlorophyll and carotenoid concentrations [40]. By the final stages of maturation, the fruit typically exhibits a violet or purple coloration due to anthocyanin accumulation [41]. Therefore, the spectral signature differences observed at the extremes of the parameter distributions may reflect

pigment-related coloration changes accompanying the dynamics of these quality indicators during maturation.

In contrast, the inconsistent spectral trend observed for TA may be explained by its weaker link to ripening dynamics. TA is more strongly influenced by fruit damage caused by environmental factors or biotic stressors than by physiological ripening processes [8]. For instance, TA can increase during early ripening due to physical or pathological fruit damage, disrupting its correlation with the pigment-related coloration changes associated with ripening.

Furthermore, the spectral range of the proposed sensor (410–940 nm) may represent a physical limitation for TA assessment. The chemical bonds responsible for TA, mainly C=O and specific O-H stretching vibrations of the carboxyl groups in organic acids, have their primary absorption features in the short-wave infrared (SWIR) and mid-infrared (MIR) regions, well beyond 1000 nm [34,42]. Since the device operates within the VIS and very short NIR range, it is unlikely to directly detect these functional groups, and any potential weak overtones would likely be masked by the dominant water absorption bands. Therefore, TA prediction appears to rely heavily on the aforementioned indirect correlation with pigment changes during physiological ripening. When fruit damage alters TA independently of the ripening stage, this secondary link may be disrupted, potentially leading to the poor predictive performance observed.

On the other hand, the device's coverage of the NIR spectrum (810–940 nm) captures regions associated with the first and second overtones of O–H bond stretching vibrations predominantly belonging to water molecules [42]. As such, the spectral variations observed in these bands in relation to OCFW, OCDM, and M levels may also be attributable to differences in water content within the fruit tissues.

One significant advantage of ANN methods is that the weights of the model's parameters are automatically adjusted during training. Consequently, a precise a priori knowledge of how every single spectral band relates to the target parameters is not required. For this reason, and considering the limited number of spectral bands acquired by the sensor (18), none of them was discarded for feeding the ANN models.

To evaluate the reliability of the proposed methodology, an external validation was carried out using unseen samples. The results revealed promising relationships between the models' response and the reference chemical values for OCFW ($R^2p = 0.86$), OCDM ($R^2p = 0.86$), and M ($R^2p = 0.89$). Statistically, these elevated coefficients of determination indicate that a high percentage of the variance contained in the target variables is accounted for by the models' response, demonstrating a strong capability to model the underlying trends [34]. However, while R^2 assesses correlation, the accuracy of the estimation is better represented by the Root Mean Square Error of Prediction (RMSEP), which quantifies the total magnitude of the error. The obtained RMSEP values (1.29% for OCFW, 2.74% for OCDM, and 1.66% for M) indicate the average deviation of the predictions from the true values. When normalized to the scale of the data (rRMSEP), these errors represented 6.78%, 6.25%, and 2.88% of the mean values, respectively, confirming that the moisture model (M) achieved the highest accuracy, followed by the oil content models.

In terms of precision, the Standard Error of Prediction (SEP) isolated the random component of the error, disregarding systematic bias. The SEP values (1.30% for OCFW, 2.52% for OCDM, 1.65% for M) were very close to the RMSEP values, which suggests that the systematic bias in the external validation was negligible. The relative precision (rSEP) stood at 6.85%, 5.75%, and 2.86%, respectively. Conversely, the model aimed to estimate TA exhibited a notably lower performance ($R^2p = 0.21$), with a significantly higher relative error (rRMSEP = 42.38%) and dispersion (rSEP = 40.21%), indicating that the spectral information acquired was insufficient to model this parameter accurately.

The robustness of the models was verified by comparing the calibration and validation statistics. The metrics obtained during the training phase (R^2c, RMSEC, rRMSEC, SEC, and rSEC) were remarkably similar to those observed in the external validation (Tables 2 and 3). For instance, the gap between RMSEC (accuracy in training) and RMSEP (accuracy in validation) was minimal for the effective models. These small differences were within expected margins, considering the different sizes of both datasets. This consistency confirms that the training conditions were suitable and allows for ruling out overfitting, demonstrating that the ANNs successfully generalized the spectral features rather than memorizing the training data.

Finally, the Ratio of Performance to Deviation (RPD) provided a comprehensive measure of the models' utility by relating prediction accuracy to the natural variability of the sample set. The RPD values obtained (2.38 for OCFW, 2.47 for OCDM, and 2.58 for M) surpass the threshold of 2.0, which is generally considered indicative of good predictive potential suitable for approximate quantitative predictions and screening [35,36]. These results align with the low rRMSEP and high R^2 values, supporting the reliability of the sensor for monitoring oil and moisture. In contrast, the low RPD value obtained for TA (1.12) confirms the limited predictive capability for this parameter.

The spectral range acquired by the proposed device (410–940 nm) encompasses mainly the VIS domain (410–760 nm). Therefore, part of the theoretical basis for the strong performance shown by the ANN models in estimating OCFW, OCDM, and M may be attributed to the correlation between fruit coloration changes and the dynamics of these parameters. This assumption aligns with the fact that fruit colour is commonly used as a visual indicator of maturity level, often expressed through a maturity index [6,43]. This correlation could explain the poor performance of the models developed to estimate TA, as TA is highly influenced by fruit damage rather than ripening dynamics [8]. When comparing the spectral signatures at the extremes of the TA histograms, only minor differences were observed. Thus, damage sustained by the fruit during early ripening stages could elevate TA levels, weakening its correlation with colour changes and, in turn, reducing the quality of the input features used by the ANN model. On the other hand, the inclusion of spectral bands in the NIR domain (810–940 nm), which have been associated with water absorption features [42], could account for the better results achieved by the model aimed at estimating M.

These results are particularly striking when compared to those reported by Noguera et al. 2022 [31], in which a system based on the same AS7265x spectral sensor was evaluated under laboratory conditions to predict the same quality parameters. In that study, TA estimation models achieved the highest performance (R^2 = 0.86), followed by M (R^2 = 0.78) and OCFW (R^2 = 0.62). However, there are significant differences in both device configuration and experimental conditions that may account for these disparities.

Firstly, the device used in the present study was adapted for portability, which required major hardware modifications, including new components (e.g., a different light source, light diffusion film) and a different interaction geometry between the light source, sample, and sensor. Secondly, the dataset in the 2022 study included four olive varieties collected at their optimal harvest stage according to growers' criteria, whereas in the present work, periodic sampling was performed throughout the entire ripening process of a single variety. Thirdly, the reference methods used for sample characterization differed. Lastly, spectral measurements in this study were acquired in the field rather than under controlled laboratory conditions. All of these factors hinder a direct comparison and likely explain the observed differences in model performance.

The use of spectral sensing to characterize intact olive fruits is a relatively mature topic [44]. There have been numerous studies evaluating different spectral sensors and chemometric techniques to assess chemical parameters of olives in a non-

destructive way. However, most of them have been accomplished under laboratory conditions [4,17–22,31,39,44–49]. For example, Salguero-Chaparro et al. [4] conducted a study to assess moisture and fat content in intact olive fruits from 50 different varieties using a NIR instrument operating in the 400–2500 nm range. Their PLS models achieved RPD values of 2.32 for moisture, 2.08 for fat content, and 1.70 for acidity. On the other hand, Grassi et al. [39] compared the performance of a Fourier transform (FT)-NIR spectrometer, equipped with both an integrating sphere and a fiber optic probe, and a Vis/NIR handheld device for predicting oil content and moisture of intact olive fruits. They used PLS as a modelling tool, obtaining determination coefficients of prediction ranging from 0.77 to 0.84 for moisture and from 0.64 to 0.78 for oil content. The best results for both quality parameters were achieved with the FT-NIR spectrometer equipped with the fiber optic probe.

The adaptation of these methodologies to in-field operations involves several advantages, such as simultaneous multiple measurements and real-time decision making. However, the effect of environmental variables (temperature, moisture, pressure, ambient light, etc.) on NIR measurements, which are easily controlled under laboratory conditions, makes it difficult to apply such methodologies in the field. In this sense, Garcia and León [23] developed PLSR models using spectral data acquired under field conditions with a portable AOTF-NIR spectrophotometer to estimate oil content and moisture in intact olive fruits. Their models achieved R^2 values of 0.85 for oil content and 0.83 for moisture. The authors compared the performance of models trained with field-acquired data against those built using laboratory-acquired data, concluding that, although slightly better results were obtained under controlled laboratory conditions, the models developed with field data also showed sufficient accuracy for practical applications.

The results reported in the present study are comparable to those observed in previous investigations. Beyond the specific device used for spectral data acquisition, one of the main differences in the proposed methodology lies in the application of ANN as a modelling tool. Traditionally, PLSR has been the most commonly employed technique in this context. PLSR intrinsically addresses the issue of collinearity among input variables by transforming them into uncorrelated principal components through data compression. As such, it is particularly effective when the number of explanatory variables is relatively high compared to the number of observations, a typical scenario when using NIR spectrometers [50]. In contrast, the device proposed in this study is based on a multispectral sensor with a capture range limited to only 18 spectral bands. Consequently, the lower dimensionality of the input data enables the ANN to efficiently adjust the feature space during training through automatic weight optimization. Furthermore, the non-linear nature of ANN provides high flexibility to model the complex underlying relationships between input variables and response parameters, using local or specific equations. It is important to note that the selection of ANN was based on its suitability for low-dimensional, non-linear data rather than an absolute superiority over traditional linear methods like PLSR, which remain highly effective for high-dimensional full-spectrum data. The results demonstrate that ANN is a robust and suitable modeling tool for this specific low-cost sensor configuration. The capability of ANN models to capture the intricate non-linear relationships that often exist between biophysical vegetation parameters and spectral data has been previously reported in the literature [51–57].

An advantage of the proposed methodology is its accessibility in terms of cost and ease of use. NIR spectrometers have a higher spectral resolution than the one used in this work, which allows to register multiple narrow spectral bands further from the VIS domain. Therefore, the acquired spectral signature is more descriptive than those obtained with multispectral devices like the one used in this research. The mentioned features of these systems generally involve prohibitive costs for practical industrial use. Further, the

complexity of these devices requires expert personnel for their operation. For these reasons, these methodologies have limited applications for smallholder farming settings, which represent a significant fraction of the productive sector in many countries. Thus, it is important to develop user-friendly and affordable innovative solutions for all kinds of producers. The present study aims to contribute toward this objective, as the models developed herein could potentially be implemented in the device's embedded software, enabling it to autonomously estimate the modeled quality parameters during future practical applications. In this sense, the present research follows a trend that has occurred in recent decades, which has been motivated by the improvements of the electronics industry [42,58,59]. Enhanced hardware technology has provided sensors delivering higher performance at lower prices. Furthermore, innovative software designs offer algorithms that enhance the capacity to model biophysical parameters of vegetation by means of the information acquired by sensors. These two concepts are exemplified in the present research.

Despite these promising aspects, it is essential to acknowledge the inherent technical limitations of the low-cost multispectral sensor utilized in this study. First, as previously discussed, the restricted band range (410–940 nm) lacks access to the SWIR and MIR regions, limiting the direct assessment of specific chemical parameters like TA. Second, compared to standard high-end laboratory spectrometers, low-cost optical components generally exhibit a lower signal-to-noise ratio (SNR) and reduced detection sensitivity. These constraints can influence the precision of the acquired spectral signatures and may explain a portion of the residual errors observed in the estimation models. Acknowledging these limitations is crucial for defining the scope of application of the proposed device, while future iterations of this technology could explore hardware optimizations to improve SNR and sensitivity without compromising affordability.

5. Conclusions

This research presents the evaluation of a low-cost multispectral device specifically designed to quantitatively assess the quality status of intact olive fruits under field conditions. The obtained results indicate a high potential for estimating OCFW, OCDM, and M, proving suitable even for practical applications. In the case of TA, the results were unsatisfactory, indicating a lack of correlation between the information acquired by the sensor and this parameter.

The successful estimation of OCFW, OCDM, and M paves the way for the implementation of an olive quality appraisal system affordable for growers of all scales. Moreover, the low cost of the sensor itself facilitates its integration into wireless sensor networks or robotic devices. The availability of a cost-effective and user-friendly methodology for assessing olive quality would benefit the entire olive sector. The improvement of the spatial-temporal resolution in monitoring, compared to traditional chemical methods, offers a more accurate and complete view of the ripening process. This would allow growers to apply precision farming techniques to manage the heterogeneity regarding ripening rates across different field plots. Furthermore, it would enable adapting the harvest time to this variability, thereby diversifying production to obtain fruit with specific characteristics. This would provide the olive oil industry with better resources to develop products with greater added value. Additionally, by expanding accessibility to this kind of technology, the understanding of olive maturation and how biotic and abiotic factors influence it will be enhanced.

However, it is important to note that the methodology proposed in this paper requires further research before it can be transferred to a commercial scale. Increasing the sample size over additional crop years could improve the robustness of the predictive models. Furthermore, it is necessary to test the models with other olive varieties to verify their

generalization capability. Therefore, there is still room for improving the results presented in this study in future research.

Author Contributions: Conceptualization, M.N., B.M., A.A. and J.M.A.; methodology, M.N., B.M., A.A. and J.M.A.; software, B.M.; validation, M.N.; formal analysis, M.N.; investigation, M.N.; resources, B.M. and A.A.; data curation, M.N., B.M. and A.A.; writing—original draft preparation, M.N. and B.M.; writing—review and editing, A.A. and J.M.A.; visualization, M.N.; supervision, J.M.A.; project administration, B.M. and A.A.; funding acquisition, B.M. and A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Research Project PID2020-119217RA-I00 funded by MCIN/AEI/10.13039/501100011033; grant IJC2019-040114-I funded by MCIN/AEI/10.13039/501100011033; grant CNS2022_136137 funded by MICIU/AEI/10.13039/501100011033 and by "European Union NextGenerationEU/PRTR"; the Interreg Cooperation Program VI-A SPAIN-PORTUGAL (POCTEP) 2021-27 and co-financed with ERDF, grant number 0067_OLIVARIA_5_E, within the scope of the OlivarIA Project.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to commercial confidentiality restrictions.

Acknowledgments: The authors acknowledge Francisco Dominguez Calvo, the Nuestra Señora de la Oliva manager, for providing the experimental field in which the study was conducted, as well as the technical support. The authors also acknowledge Kimoto LDT for supplying the material used for making the light-diffuser film. The authors utilized the Gemini Pro 1.5 model for language polishing and editing of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

NMR	Nuclear magnetic resonance
VIS	Visible
NIR	Near infrared
AOFT	Acousto-optic tunable filter
PLSR	Partial least squares regression
ANN	Artificial neural network
PCB	Printed circuit board
3D	Three-dimensional
FWHM	Full width at half maximum
IR	Infrared
IDE	Integrated development environment
LED	Light-emitting diode
PC	Personal computer
PLA	Polylactic acid
ANOVA	Analysis of variance
OC	Oil content
M	Moisture
TA	Titratable acidity
OCDM	Oil content per dry matter
OCFW	Oil content per fresh weight
RS	Random sampling

R^2c	Coefficient of determination for calibration
R^2p	Coefficient of determination for external validation
RMSEC	Root mean square error of calibration
RMSEP	Root mean square error of prediction
SEC	Standard error of calibration
SEP	Standard error of prediction
RPD	Ratio of performance to deviation
rRMSEC	Coefficient of variation in the RMSEC
rRMSEP	Coefficient of variation in the RMSEP
rSEC	Coefficient of variation in the SEC
rSEP	Coefficient of variation in the SEP
FT	Fourier transform

References

1. Aybar, V.E.; de Melo-Abreu, J.P.; Searles, P.S.; Matias, A.C.; del Río, C.; Caballero, J.M.; Rousseaux, M.C. Evaluation of Olive Flowering at Low Latitude Sites in Argentina Using a Chilling Requirement Model. *Span. J. Agric. Res.* **2015**, *13*, e0901. [CrossRef]
2. Rallo, L.; Caruso, T.; Díez, C.M.; Campisi, G. Olive Growing in a Time of Change: From Empiricism to Genomics. In *The Olive Tree Genome*; Rugini, E., Baldoni, L., Muleo, R., Sebastiani, L., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 55–64.
3. Fernández-Espinosa, A.J. Combining PLS Regression with Portable NIR Spectroscopy to On-Line Monitor Quality Parameters in Intact Olives for Determining Optimal Harvesting Time. *Talanta* **2016**, *148*, 216–228. [CrossRef] [PubMed]
4. Salguero-Chaparro, L.; Baeten, V.; Fernández-Pierna, J.A.; Peña-Rodríguez, F. Near Infrared Spectroscopy (NIRS) for on-Line Determination of Quality Parameters in Intact Olives. *Food Chem.* **2013**, *139*, 1121–1126. [CrossRef] [PubMed]
5. Hermoso, M.; González, J.; Uceda, M.; García-Ortiz, A.; Morales, J.; Frías, L.; Fernández, A. *Elaboración de Aceite de Oliva de Calidad, Obtención Por El Sistema de Dos Fases*, 3rd ed.; Junta de Andalucia—Consejería de Agricultura y Pesca: Sevilla, Spain, 1996.
6. Uceda, M.; Frias, L. Harvest Dates. Evolution of the Fruit Oil Content, Oil Composition and Oil Quality. In Proceedings of the II Seminario Oleícola Internacional, International Olive Oil Council, Cordoba, Spain, 6–17 October 1975; pp. 125–128.
7. Fernández-Escobar, R.; Braz Frade, R.; Lopez Campayo, M.; Beltrán Maza, G. Effect of Nitrogen Fertilization on Fruit Maturation of Olive Trees. *Acta Hortic.* **2014**, *1057*, 101–105. [CrossRef]
8. Barranco Navero, D.; Fernandez-Escobar, R.; Rallo Romero, L. *El Cultivo del Olivo*, 7th ed.; Fernandez Escobar, R., Rallo Romero, L., Barranco Navero, D., Eds.; Mundi-Prensa Libros: Madrid, Spain, 2017.
9. Cecchi, L.; Migliorini, M.; Mulinacci, N. Virgin Olive Oil Volatile Compounds: Composition, Sensory Characteristics, Analytical Approaches, Quality Control, and Authentication. *J. Agric. Food Chem.* **2021**, *69*, 2013–2040. [CrossRef] [PubMed]
10. *ISO 660:2020*; Animal and Vegetable Fats and Oils–Determination of Acid Value and Acidity. International Organization of Standardization: Geneva, Switzerland, 2020.
11. *UNE 55030:1961*; Determination of the Content in Total Fat of Olives. Asociación Española de Normalización y Certificación: Madrid, Spain, 1961.
12. *ISO 662:2016*; Animal and Vegetable Fats and Oils–Determination of Moisture and Volatile Matter Content. International Organization of Standardization: Madrid, Spain, 2016.
13. Cattaneo, T.M.P.; Stellari, A. Review: NIR Spectroscopy as a Suitable Tool for the Investigation of the Horticultural Field. *Agronomy* **2019**, *9*, 503. [CrossRef]
14. García Sánchez, P.A.; Ramos Martos, N.; Ballesteros, E. 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 Aceites* **2005**, *56*, 220–227. [CrossRef]
15. Jiménez, A.; Izquierdo, E.; Rodríguez, F.; Dueñas, J.I.; Tortosa, C. Determination of Fat and Moisture in Olives by Nearinfrarred Reflectance Spectroscopy. *Grasas Aceites* **2000**, *51*, 311–315. [CrossRef]
16. Bendini, A.; Cerretani, L.; Di Virgilio, F.; Belloni, P.; Lercker, G.; Toschi, T.G. In-Process Monitoring in Industrial Olive Mill by Means of FT-NIR. *Eur. J. Lipid Sci. Technol.* **2007**, *109*, 498–504. [CrossRef]
17. Cayuela, J.A.; Camino, M.D.C.P. Prediction of Quality of Intact Olives by near Infrared Spectroscopy. *Eur. J. Lipid Sci. Technol.* **2010**, *112*, 1209–1217. [CrossRef]
18. Kavdir, I.; Buyukcan, M.B.; Lu, R.; Kocabiyik, H.; Seker, M. Prediction of Olive Quality Using FT-NIR Spectroscopy in Reflectance and Transmittance Modes. *Biosyst. Eng.* **2009**, *103*, 304–312. [CrossRef]
19. Beghi, R.; Giovenzana, V.; Civelli, R.; Cini, E.; Guidetti, R. Characterisation of Olive Fruit for the Milling Process by Using Visible/near Infrared Spectroscopy. *J. Agric. Eng.* **2013**, *44*, 8. [CrossRef]

20. Giovenzana, V.; Beghi, R.; Romaniello, R.; Tamborrino, A.; Guidetti, R.; Leone, A. Use of Visible and near Infrared Spectroscopy with a View to On-Line Evaluation of Oil Content during Olive Processing. *Biosyst. Eng.* **2018**, *172*, 102–109. [CrossRef]
21. Bellincontro, A.; Taticchi, A.; Servili, M.; Esposto, S.; Farinelli, D.; Mencarelli, F. Feasible Application of a Portable NIR-AOTF Tool for on-Field Prediction of Phenolic Compounds during the Ripening of Olives for Oil Production. *J. Agric. Food Chem.* **2012**, *60*, 2665–2673. [CrossRef] [PubMed]
22. Bellincontro, A.; Caruso, G.; Mencarelli, F.; Gucci, R. Oil Accumulation in Intact Olive Fruits Measured by near Infrared Spectroscopy-Acousto-Optically Tunable Filter. *J. Sci. Food Agric.* **2013**, *93*, 1259–1265. [CrossRef] [PubMed]
23. Gracia, A.; León, L. Non-Destructive Assessment of Olive Fruit Ripening by Portable near Infrared Spectroscopy. *Grasas Aceites* **2011**, *62*, 268–274. [CrossRef]
24. Trang, N.M.; Duy, T.K.; Huyen, T.T.N.; Danh, L.V.Q.; Dinh, A. An Investigation into the Use of a Low-Cost NIR Integrated Circuit Spectrometer to Measure Chlorophyll Content Index. *Int. J. Innov. Technol. Explor. Eng.* **2019**, *8*, 35–38.
25. Li, M.; Qian, Z.; Shi, B.; Medlicott, J.; East, A. Evaluating the Performance of a Consumer Scale SCiO^{TM} Molecular Sensor to Predict Quality of Horticultural Products. *Postharvest Biol. Technol.* **2018**, *145*, 183–192. [CrossRef]
26. Kitić, G.; Tagarakis, A.; Cselyuszka, N.; Panić, M.; Birgermajer, S.; Sakulski, D.; Matović, J. A New Low-Cost Portable Multispectral Optical Device for Precise Plant Status Assessment. *Comput. Electron. Agric.* **2019**, *162*, 300–308. [CrossRef]
27. Moinard, S.; Brunel, G.; Ducanchez, A.; Crestey, T.; Rousseau, J.; Tisseyre, B. Testing the Potential of a New Low-Cost Multispectral Sensor for Decision Support in Agriculture. *Precis. Agric.* **2021**, *21*, 411–418. [CrossRef] [PubMed]
28. Leon-Salas, W.D.; Rajendran, J.; Vizcardo, M.A.; Postigo-Malaga, M. Measuring Photosynthetically Active Radiation with a Multi-Channel Integrated Spectral Sensor. In Proceedings of the 2021 IEEE International Symposium on Circuits and Systems (ISCAS), Daegu, Republic of Korea, 21–28 May 2021; pp. 1–5.
29. Stevens, J.D.; Murray, D.; Sozzi, M.; Cogato, A.; Laroche-Pinel, E.; Stevens, J.D.; Murray, D.; Diepeveen, D.; Toohey, D. Development and Testing of an IoT Spectroscopic Nutrient Monitoring System for Use in Micro Indoor Smart Hydroponics. *Horticulturae* **2023**, *9*, 185. [CrossRef]
30. Lopin, P.; Nawsang, P.; Laywisadkul, S.; Lopin, K.V. Evaluation of Low-Cost Multi-Spectral Sensors for Measuring Chlorophyll Levels Across Diverse Leaf Types. *Sensors* **2025**, *25*, 2198. [CrossRef] [PubMed]
31. Noguera, M.; Millan, B.; Aquino, A.; Andújar, J.M. Methodology for Olive Fruit Quality Assessment by Means of a Low-Cost Multispectral Device. *Agronomy* **2022**, *12*, 979. [CrossRef]
32. Noguera, M.; Millan, B.; Andújar, J.M. New, Low-Cost, Hand-Held Multispectral Device for In-Field Fruit-Ripening Assessment. *Agriculture* **2022**, *13*, 4. [CrossRef]
33. Savitzky, A.; Golay, M.J.E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **1964**, *36*, 1627–1639. [CrossRef]
34. Williams, P.; Manley, M.; Antoniszyn, J. *Near Infrared Technology: Getting the Best Out of Light*; African Sun Media: Stellenbosch, South Africa, 2019.
35. Paz, M.; Fern, J.; Rom, S.M.; Cebri, C. Use of the Hyperspectral Imaging to Estimate the Volatile Composition of Tempranillo Grape Berries during Ripening. *Sci. Hortic.* **2024**, *337*, 3537. [CrossRef]
36. Nicolaï, B.M.; Beullens, K.; Bobelyn, E.; Peirs, A.; Saeys, W.; Theron, K.I.; Lammertyn, J. Nondestructive Measurement of Fruit and Vegetable Quality by Means of NIR Spectroscopy: A Review. *Postharvest Biol. Technol.* **2007**, *46*, 99–118. [CrossRef]
37. *COI/T.15/NC No. 3*; Trade Standard Applying to Olive Oils and Olive-Pomace Oils. International Olive Council: Madrid, Spain, 2022.
38. European Commission. Comisión Europea Reglamento Delegado (UE) 2022/1288 de La Comisión de 6 de Abril de 2022 Por El Que Se Completa El Reglamento (UE) 2019/2088 Del Parlamento Europeo y Del Consejo. *Diario Oficial de la Unión Europea* **2022**, *L 196*, 1–72. [CrossRef]
39. Grassi, S.; Jolayemi, O.S.; Giovenzana, V.; Tugnolo, A.; Squeo, G.; Conte, P.; De Bruno, A.; Flamminii, F.; Casiraghi, E.; Alamprese, C. Near Infrared Spectroscopy as a Green Technology for the Quality Prediction of Intact Olives. *Foods* **2021**, *10*, 1042. [CrossRef] [PubMed]
40. Salvador, M.D.; Aranda, F.; Fregapane, G. Influence of Fruit Ripening on 'Cornicabra' Virgin Olive Oil Quality A Study of Four Successive Crop Seasons. *Food Chem.* **2001**, *73*, 45–53. [CrossRef]
41. Roca, M.; Mínguez-Mosquera, M.I. Change in the Natural Ratio between Chlorophylls and Carotenoids in Olive Fruit during Processing for Virgin Olive Oil. *JAOCS J. Am. Oil Chem. Soc.* **2001**, *78*, 133–138. [CrossRef]
42. Walsh, K.B.; Blasco, J.; Zude-Sasse, M.; Sun, X. Visible-NIR 'Point' Spectroscopy in Postharvest Fruit and Vegetable Assessment: The Science behind Three Decades of Commercial Use. *Postharvest Biol. Technol.* **2020**, *168*, 111246. [CrossRef]
43. Dag, A.; Kerem, Z.; Yogev, N.; Zipori, I.; Lavee, S.; Ben-David, E. Influence of Time of Harvest and Maturity Index on Olive Oil Yield and Quality. *Sci. Hortic.* **2011**, *127*, 358–366. [CrossRef]
44. León, L.; Rallo, L.; Garrido, A. Análisis de Aceituna Intacta Mediante Espectroscopia En El Infrarrojo Cercano (NIRS): Una Herramienta de Utilidad En Programas de Mejora de Olivo. *Grasas Aceites* **2003**, *54*, 41–47. [CrossRef]

45. Salguero-Chaparro, L.; Peña-Rodríguez, F. On-Line versus off-Line NIRS Analysis of Intact Olives. *LWT Food Sci. Technol.* **2014**, *56*, 363–369. [CrossRef]
46. Cayuela, J.A.; García, J.M.; Caliani, N. NIR Prediction of Fruit Moisture, Free Acidity and Oil Content in Intact Olives. *Grasas Aceites* **2009**, *60*, 194–202. [CrossRef]
47. Saha, U.; Jackson, D. Analysis of Moisture, Oil, and Fatty Acid Composition of Olives by near-Infrared Spectroscopy: Development and Validation Calibration Models. *J. Sci. Food Agric.* **2018**, *98*, 1821–1831. [CrossRef] [PubMed]
48. Cirilli, M.; Bellincontro, A.; Urbani, S.; Servili, M.; Esposto, S.; Mencarelli, F.; Muleo, R. On-Field Monitoring of Fruit Ripening Evolution and Quality Parameters in Olive Mutants Using a Portable NIR-AOTF Device. *Food Chem.* **2016**, *199*, 96–104. [CrossRef] [PubMed]
49. Giovenzana, V.; Beghi, R.; Civelli, R.; Trapani, S.; Migliorini, M.; Cini, E.; Zanoni, B.; Guidetti, R. Rapid Determination of Crucial Parameters for the Optimization of Milling Process by Using Visible/near Infrared Spectroscopy on Intact Olives and Olive Paste. *Ital. J. Food Sci.* **2017**, *29*, 357–369. [CrossRef]
50. Verrelst, J.; Malenovský, Z.; Van der Tol, C.; Camps-Valls, G.; Gastellu-Etchegorry, J.-P.; Lewis, P.; North, P.; Moreno, J. Quantifying Vegetation Biophysical Variables from Imaging Spectroscopy Data: A Review on Retrieval Methods. *Surv. Geophys.* **2019**, *40*, 589–629. [CrossRef] [PubMed]
51. Allouche, Y.; López, E.F.; Maza, G.B.; Márquez, A.J. Near Infrared Spectroscopy and Artificial Neural Network to Characterise Olive Fruit and Oil Online for Process Optimisation. *J. Near Infrared Spectrosc.* **2015**, *23*, 111–121. [CrossRef]
52. Kalacska, M.; Lalonde, M.; Moore, T.R. Estimation of Foliar Chlorophyll and Nitrogen Content in an Ombrotrophic Bog from Hyperspectral Data: Scaling from Leaf to Image. *Remote Sens. Environ.* **2015**, *169*, 270–279. [CrossRef]
53. Wang, F.; Huang, J.; Wang, Y.; Liu, Z.; Peng, D.; Cao, F. Monitoring Nitrogen Concentration of Oilseed Rape from Hyperspectral Data Using Radial Basis Function. *Int. J. Digit. Earth* **2013**, *6*, 550–562. [CrossRef]
54. Huang, Z.; Turner, B.J.; Dury, S.J.; Wallis, I.R.; Foley, W.J. Estimating Foliage Nitrogen Concentration from HYMAP Data Using Continuum Removal Analysis. *Remote Sens. Environ.* **2004**, *93*, 18–29. [CrossRef]
55. Fazari, A.; Pellicer-Valero, O.J.; Gómez-Sanchıs, J.; Bernardi, B.; Cubero, S.; Benalia, S.; Zimbalatti, G.; Blasco, J. Application of Deep Convolutional Neural Networks for the Detection of Anthracnose in Olives Using VIS/NIR Hyperspectral Images. *Comput. Electron. Agric.* **2021**, *187*, 106252. [CrossRef]
56. Pôças, I.; Gonçalves, J.; Costa, P.M.; Gonçalves, I.; Pereira, L.S.; Cunha, M. Hyperspectral-Based Predictive Modelling of Grapevine Water Status in the Portuguese Douro Wine Region. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *58*, 177–190. [CrossRef]
57. Mayr, S.; Beć, K.B.; Grabska, J.; Wiedemair, V.; Pürgy, V.; Popp, M.A.; Bonn, G.K.; Huck, C.W. Challenging Handheld NIR Spectrometers with Moisture Analysis in Plant Matrices: Performance of PLSR vs. GPR vs. ANN Modelling. *Spectrochim. Acta-Part A Mol. Biomol. Spectrosc.* **2021**, *249*, 9342. [CrossRef] [PubMed]
58. Bec, K.B.; Grabska, J.; Huck, C.W. Miniaturized NIR Spectroscopy in Food Analysis and Quality Control: Promises, Challenges, and Perspectives. *Foods* **2022**, *11*, 1465. [CrossRef] [PubMed]
59. Krause, J.; Grüger, H.; Gebauer, L.; Zheng, X.; Knobbe, J.; Pügner, T.; Kicherer, A.; Gruna, R.; Längle, T.; Beyerer, J. SmartSpectrometer—Embedded Optical Spectroscopy for Applications in Agriculture and Industry. *Sensors* **2021**, *21*, 4476. [CrossRef] [PubMed]