




## Article

# Quality Evaluation of Fair-Trade Cocoa Beans from Different Origins Using Portable Near-Infrared Spectroscopy (NIRS)

Matteo Forte <sup>1</sup>, Sarah Currò <sup>2,\*</sup> , Davy Van de Walle <sup>3</sup>, Koen Dewettinck <sup>3</sup>, Massimo Mirisola <sup>4</sup>, Luca Fasolato <sup>2</sup>   
and Paolo Carletti <sup>1</sup> 

- <sup>1</sup> Dipartimento di Agronomia, Animali, Alimenti, Risorse Naturali e Ambiente, Università degli Studi di Padova, Viale dell'Università 16, 35020 Legnaro, PD, Italy
- <sup>2</sup> Dipartimento di Biomedicina Comparata e Alimentazione, Università degli Studi di Padova, Viale dell'Università 16, 35020 Legnaro, PD, Italy
- <sup>3</sup> Food Structure and Function Research Group, Department of Food Technology, Safety and Health, Faculty of Bioscience Engineering, Ghent University 653, 9000 Ghent, Belgium
- <sup>4</sup> Dipartimento di Medicina Animale, Produzioni e Salute, Università degli Studi di Padova, Viale dell'Università 16, 35020 Legnaro, PD, Italy
- \* Correspondence: sarah.curro@unipd.it

**Abstract:** Determining cocoa bean quality is crucial for many players in the international supply chain. However, actual methods rely on a cut test protocol, which is limited by its subjective nature, or on time-consuming, expensive and destructive wet-chemistry laboratory procedures. In this context, the application of near infrared (NIR) spectroscopy, particularly with the recent developments of portable NIR spectrometers, may represent a valuable solution for providing a cocoa beans' quality profile, in a rapid, non-destructive, and reliable way. Monitored parameters in this work were dry matter (DM), ash, shell, fat, protein, total polyphenols, fermentation index (FI), titratable acidity (TA) and pH. Different chemometric analyses were performed on the spectral data and calibration models were developed using modified partial least squares regression. Prediction equations were validated using a fivefold cross-validation and a comparison between the different prediction performances for the portable and benchtop NIR spectrometers was provided. The NIRS benchtop instrument provided better performance of quantification considering the whole than the portable device, showing excellent prediction capability in protein and DM quantification. On the other hand, the NIRS portable device, although showing lower but valuable performance of prediction, can represent an appealing alternative to benchtop instruments for food business operators, being applicable in the field.

**Keywords:** *Theobroma cacao* L.; dry matter; chemometrics; fermentation index; protein content



**Citation:** Forte, M.; Currò, S.; Van de Walle, D.; Dewettinck, K.; Mirisola, M.; Fasolato, L.; Carletti, P. Quality Evaluation of Fair-Trade Cocoa Beans from Different Origins Using Portable Near-Infrared Spectroscopy (NIRS). *Foods* **2023**, *12*, 4. <https://doi.org/10.3390/foods12010004>

Academic Editors: Mourad Kharbach and Samuli Urpelainen

Received: 23 November 2022

Revised: 15 December 2022

Accepted: 16 December 2022

Published: 20 December 2022



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

With an ancient history starting in the Preclassic period (1200–400 B.C.) with consumption among the Olmec and other pre-Colombian populations of the Americas [1], cocoa is now a ubiquitous food. Cocoa appears today in many different forms, mainly in chocolate, with consumption averaging around 8 kg per person per annum in many European countries [2]. The top four countries, which account for nearly 65% of the total world chocolate production, are the USA, Germany, Switzerland, and Belgium, which in terms of retail sales reach respectively USD 20, USD 10, USD 14, and USD 12 billion per year [3]. The consumption of cocoa and cocoa-based products is of great interest both for the highly appreciated sensorial profile and for the possible beneficial health effects which are being studied in current times [4,5].

Cocoa beans represent the essential raw material for chocolate, and they are obtained from the *Theobroma cacao* L. tree, which is almost exclusively cultivated in tropical/developing countries, where it represents a source of export earnings both at the level

of families, communities, and nations [6]. Africa covers more than 75% of the world's total cocoa beans production and Côte d'Ivoire is the largest cocoa bean exporter in the world, with more than USD four billion exported in 2020, followed by Ghana and Ecuador.

Once the cacao pods have been harvested, they must undergo post-harvest operations on farms and plantations before becoming the so-called cocoa beans, which are then traded in the international market and processed into final industrial products. The post-harvest processing comprise pod opening and removal of beans from the pod, bean-pulp mass fermentation and bean drying. In this sequence, the fermentation constitutes an essential critical step for the development of desired flavor attributes of the commercial cocoa beans. In the further processing, cocoa beans are roasted, cracked and ground to give a powdery mass from which fat is expressed [7] and the release of fat ultimately leads to a liquid-like ingredient, namely cocoa liquor. Additionally, many processes are implemented in the chocolate industry that originate several products with different forms and functionalities.

Fermentation and drying constitute key farm(er)-based unit operations with strong influences in the final quality of cocoa beans and subsequent products [8].

Since 2010, approximately 4 million tons of cocoa beans have been produced annually around the world [9], and the three biggest importers are the Netherlands, Germany, and USA, with product worth 2,375,923, 1,209,366 and 1,026,931 USD imported in 2020, respectively (Sources: ITC calculations based on UN COMTRADE and ITC statistics).

The socio-economic importance and international interest towards cocoa beans are intelligible as an estimated five million farming households depend on cacao as a cash crop, and 70 per cent of cocoa is produced by smallholders living on less than USD 2 per day and relying on cocoa for 60 to 90 per cent of their income [9].

Cocoa beans are divided worldwide between “fine or flavor” (mainly Criollo and Trinitario) and “bulk” (mainly Forastero) varieties, with a common belief that fine and flavor cacao varieties receive significant price premiums in international markets. However, recent studies show how post-harvest processing has a central role in causing heterogeneity in cocoa prices, independently of the variety grown [10]. This is because although the primary factors influencing the quality attributes of cocoa beans are the cocoa tree cultivar and genotype, it is well-established that the agronomic and environmental conditions together with the harvest and post-harvest steps are crucial elements in the determination of the final quality of commercial cocoa beans [11–15].

Partially fermented or unfermented beans are prone to bitterness and astringency with poor chocolate flavor and aroma [2]. Moreover, an appropriate drying process will reduce the beans' water content of 55% to around 7% [16], preventing fermentation from continuing uncontrolled, slowing the development of molds that could give rise to unwanted, unpleasant flavors and equilibrating the beans acidity that would otherwise be excessive in the final products [17,18].

Although fermentation is considered as the “core stage” of the cocoa transformation process from seeds to chocolate, it is currently performed mostly by small third-world producers in an empirical way, with little or no technification, without control in processing conditions, originating cocoa batches of low and heterogeneous quality [13].

Chocolate and chocolate-based products sell in a very competitive market, where quality is crucial, and value is enormous. It should be clear that if the quality of cocoa beans is poor, final products will suffer this deficiency as well, and the whole industry sector will be affected [19].

Most of the existing commercial standards for cocoa beans base their quality requirements on the results of the cut test, or on sensory estimation by trained panels (Aculey et al., 2010). The cut test consists of cutting cocoa beans lengthwise, observing the number of defective beans. The ISO 1114 states that both halves of each bean shall be visually examined, and the result for each kind of defect shall be expressed as a percentage of the 300 beans examined. The ISO also defines nine categories of defects: those related to poor fermentation (slaty and violet/purple beans) and those being indicators of high FFA levels,

poor flavor and/or other contaminants (bean clusters, broken beans, smoky beans, moldy beans, germinated beans, flat beans, insect-damaged/infested beans) [20].

On the one hand, the cut test is limited by its subjective nature and does not represent a sufficiently reliable methodology for a comprehensive description of the main quality contributors [21]. On the other hand, laboratory methodologies are often demanding in terms of time and cost effectiveness, which can be critical factors, and they are mainly based on destructive determinations. Moreover, in cocoa producing countries the availability of laboratory infrastructures is poor [22].

In this context, the application of near infrared (NIR) spectroscopy, particularly with the recent developments of portable NIR spectrometers, may represent a valuable solution for providing a cocoa beans' quality profile, in a rapid, non-destructive, and reliable way. This analytical technique could be useful to both cocoa bean producers, mostly in the developing countries, and processors, mostly in the developed countries, alike.

Many researchers have already investigated quality parameters of cocoa beans through NIR spectroscopy quantifying, for instance, fat, sugars, proteins, moisture, pH and titratable acidity, polyphenols, and other volatile and non-volatile compounds [23]. Even the assessment of the authenticity of cocoa powder has been studied by identifying the country of origin of raw materials, varietal purity, or the presence of adulterants [24].

It must be noted, however, that most of the studies that successfully predicted cocoa beans' quality parameters through NIR spectroscopy mainly analyzed samples which had been purposely subjected to different degrees of fermentation, e.g., analyzing the beans at different days during the fermentation process. Doing so, the samples are not in the status in which importers/exporters normally trade them in the international market, fully fermented and dried, therefore this might not be a representative "working condition" for these actors in the cocoa supply chain. Moreover, only cocoa bean samples coming from one or few countries are most often utilized in the previous studies, hence limiting the variability that can be included in the NIR prediction model. Finally, most of the studies present in literature utilize benchtop NIR instruments, which have some notable practical disadvantages if compared with the more recent portable NIR spectrometers, which are recently being considered in the literature [25,26].

On this basis, our study aimed to predict some quality parameters of commercial cocoa bean samples using portable NIR spectrometers, also in comparison with a benchtop spectrometer, on both whole and ground samples. Cocoa beans were provided by an Italian fair-trade importer and all the samples have been produced to meet internationally accepted merchantable quality standards: well-fermented and dry. The ultimate purpose of the study was, therefore, to evaluate the possibility of using portable NIR spectrometers in commercial contexts, while also assessing the practicability of using NIR spectroscopy on whole fermented cocoa beans to rapidly predict main quality parameters.

## 2. Materials and Methods

### 2.1. Samples

Fifty-six samples from commercially available cocoa beans have been provided by Altromercato Impresa Sociale Soc. Coop. (Via 9. Crispi, Bolzano, Italy), a major importer of fair-trade products in Italy in 2021. Thirty-three of them came from Africa while the remaining twenty-three came from South America. All the samples have been produced to meet internationally accepted merchantable quality standards i.e., well-fermented and dry, free from smoky beans and abnormal or foreign odors, free from evidence of adulteration, reasonably free from living insects, virtually free from broken beans, pieces of shell, and foreign matter. African countries included Togo, Uganda, Madagascar, and Sierra Leone, while American beans came from Honduras, Ecuador, Perú, Dominican Republic, Nicaragua, and Venezuela (Supplementary Table S1). Although it was not possible to obtain accurate information on the beans' variety for each sample, documents reported that only Trinitario and Forastero varieties were utilized.

The beans were shipped in 25-ton containers which roughly corresponds to 360 bags (70 kg of cocoa beans/bag). Around 4 kg of cocoa beans from each container were sampled as representative for the 33% of the total bags' number. Of these, 200 g of cocoa beans were randomly taken for analysis and stored at  $-20\text{ }^{\circ}\text{C}$  in plastic bags until assay in a LGPv 8420 MediLine refrigerator (Liebherr, Kirchdorf an der Iller, Germany).

## 2.2. Beans Peeling and Grinding

Three random aliquots of 20 g of dry and nitrogen frozen cocoa beans from each sample were de-husked by hand. The peels and nibs were carefully collected and weighted to determine average shell percentage on a 4 digits balance (Adventurer model ARR70, OHAUS, Parsippany, NJ, USA). About 100 g of nibs (de-husked cocoa beans) were ground in a multi-purpose grinder for 45 s (3 intervals of 15 s with 10 s pause), then sifted on a 0.5 mm sieve. Before grinding, the beans were frozen with liquid nitrogen to make them brittle and avoid becoming a mash. The obtained cocoa powder for each sample was stored at  $-20\text{ }^{\circ}\text{C}$  in the dark prior to the following analyses.

## 2.3. Spectral Data Acquisition

Spectral data acquisition was performed both on whole cocoa beans and on de-husked cocoa bean powder. About 100 g of randomly chosen whole cocoa beans from each sample were scanned with a portable instrument (PoliSPEC-NIR, ITPhotonics, Breganze, Italy) and with the benchtop instrument (FOSS DS-2500 scanning monochromator FossNIR-System, Hillerød, Denmark). Both NIR data acquisitions were performed in reflectance mode, with the following parameters:

- FOSS DS-2500: scanning monochromator covering a range of 850–2500 nm at 0.5 nm intervals. Scans were performed using a slurry cup with quartz window of about a  $12.6\text{ cm}^2$  area.
- PoliSPEC-NIR: covering a range of 900–1680 nm at 2 nm intervals. Spectral data measurements were performed through a round scanning window ( $3.2\text{ cm}^2$ ) placed in direct contact with the sample surface. Each spectrum was obtained by averaging 3 data acquisitions.

## 2.4. Chemical Analyses

Unless otherwise specified, analyses were performed according to official methods of analysis (AOAC, 2016). All chemical analyses were performed in triplicate on peeled and ground cocoa beans.

### 2.4.1. Dry Matter

Dry matter is measured as subtraction of the moisture content measured using a gravimetric method based on AOAC method 931.04 [21]. Hereto, approximately 2 g of powder sample were dried at  $101\text{--}103\text{ }^{\circ}\text{C}$  to constant weight in a forced-air electric oven (UF55 Plus, Memmert, Schwabach, Germany). After the drying process was completed, the samples were immediately closed with glass lids to avoid exposure and stored in desiccators for one hour to equilibrate samples towards ambient temperature [27]. The moisture content was expressed as average percentage (%) based on loss in weight of three independent samples.

### 2.4.2. Ash

For the measurement of ashes, the sample was charred on a plate and placed in a muffle furnace (Gefran Model 1200; Gefran Spa, Brescia, Italy) at  $550\text{ }^{\circ}\text{C}$  (AOAC 972.15A). Ash content was expressed as weight percentage (%).

### 2.4.3. Fat Content

The fat content was measured by extraction with petroleum ether [21] in a TE-188 Soxhlet lipid extractor (model SOXTEC 255 Tecator-Foss Analytical, Hillerød, Denmark)

with the following parameters: 60 min boiling, 50 min washing, 15 min drying. Fat content was expressed as weight percentage (%).

#### 2.4.4. Total Protein Content

Protein determination was carried out by the Kjeldahl method, as described in AOAC 2016 (method 970.22) (model Kjeltex 2300-Foss Analytical). The protein content was calculated from the concentration of total nitrogen by applying a conversion factor of 6.25.

#### 2.4.5. Total Phenolic Content

The total phenolic content was determined according to the colorimetric method of Folin–Ciocalteu [28]. Samples were defatted using the Soxhlet method (AOAC 963.15). Defatted powder (0.05 g) was added to 10 mL of a methanol-water (70:30 *v/v*) mixture at room temperature and stirred for 45 min. After centrifugation, 0.1 mL of solution was mixed with 3 mL of distilled water and 0.5 mL Folin–Ciocalteu reagent. The mixture was stored for 3 min after which 1 mL of aqueous Na<sub>2</sub>CO<sub>3</sub> (200 g L<sup>-1</sup>) was added. The mixture was allowed to stand for 20 min at 40 °C and the total polyphenols were determined by spectrophotometry at 765 nm (spectrophotometer model Cary 60 UV-Vis Agilent Technologies Stevens Creek Blvd. Santa Clara, CA, USA). The standard curve was prepared using 0, 50, 100, 150, 200 and 250 mg L<sup>-1</sup> solutions of gallic acid in methanol. Total phenol values were expressed in terms of gallic acid equivalents (mg g<sup>-1</sup> of dry fat-free mass) [29]. The analyses were performed in triplicate.

#### 2.4.6. Fermentation Index

Fermentation index (FI) corresponds to the color change within the bean cotyledons during fermentation. This change is due to the decreasing anthocyanin content as beans progress through fermentation [30]. A 50 mg sample of previously prepared cocoa powder was weighed and mixed with 5 mL MeOH:HCl (97:3 *v/v*). Samples were extracted at 4 °C for 16–18 h, centrifuged for 5 min at 3500× *g*, and the clear supernatant was collected. Absorbance of the supernatant was read at wavelengths 460 nm and 530 nm using UV-VIS spectrophotometer (model Cary 60 UV-Vis Agilent Technologies Stevens Creek Blvd. Santa Clara, CA, USA). All the measurements were performed in triplicate.

The FI was obtained by calculating the ratio of the absorbance at 460 nm and 530 nm ( $FI = A_{460}/A_{530}$ ). Values greater than 1 are considered as well-fermented, while less than 1 as under-fermented beans [31,32]. However, it must be noted that this accounts for the Forastero variety and with some precautions for the Trinitario variety (which can contain both purple and white beans). Criollo beans do not contain anthocyanin pigments, therefore FI cannot be used to describe the fermentation level for this variety. In our study, both Trinitario and Forastero beans were used, but white beans were always absent.

#### 2.4.7. pH and Titratable Acidity

Cocoa powder (5 g) was mixed with 100 mL hot water (100 °C), stirred, and allowed to stand for 30 min. After 30 min, when the suspension was cooled up to 25 °C, it was centrifuged for 10 min at 5000 rpm and vacuum filtered through Whatman No. 4 paper filter according to AOAC 2006 methods 970.21 (pH) and 942.15 (potentiometric titration) [21], AOAC Section 42.104 (16th Ed. 1995) [21,32–35].

The pH of the filtered solution was measured with a pH-meter model PC 80 + DHS (XS Instruments, Carpi, Italy) and then 25 mL aliquots of the same solution were titrated to pH 8.1 with 0.05 M NaOH. All data were measured in triplicate. Titratable acidity results are expressed as mMol NaOH/100 g powder [34] or % acetic acid [21].

It is important to note that this procedure was not for quantifying the actual pH of the cocoa bean itself, but rather to measure the acidity derived from bean acids diffusing into water; it is useful for comparison between the pH of solutions produced by different beans [30].



### 2.5. Wavelength Selection and Chemometric Analyses

Spectral chemometric analyses were performed using firstly the wavelength selection and secondly the full spectra collected. Wavelengths selection was carried out through the interval partial least-square (iPLS) [36] and through the principal component regression (PCR) [37] by using R software version 3.2.5 (R Core Team, Auckland, New Zealand, 2016) and WinISI software (Infrasoft International, Port Matilda, PA, USA), respectively. In particular, the iPLS was carried out applying the forward mode, in which the full spectrum was subdivided in 30 intervals that are successively included in the analysis: the first step calculated 30 models (one for each interval) that were tested using the cross-validation; the interval which provides the lowest model root-mean-square error of cross-validation were selected as most informative. The selected intervals were calculated per each parameter investigated and used for the following modelling. The PCR is based on the identification of the principal factors variance among spectral absorbance data through the principal component analysis [38]. Wavelengths selection was performed on the spectra acquired with FOSS DS-2500 on cocoa powder.

The second approach considered the use of the full spectrum and mathematical treatment as reported by several authors [39–41] in foods for chemical prediction purpose. This procedure takes advantage of the mathematical treatment as multiplicative correction (MSC) of the dispersion used to correct the problems of dispersed light in reflectance spectroscopy or the spectra normalization using standard normal variation (SNV) and first or second derivatives often used to remove the deviation and slope of the baseline in the spectrum [42]. This approach was applied to spectra acquired with both instruments, on both whole and ground cocoa beans.

The calibration models were performed using the Modified PLS (MPLS) regression on wavelength selected and on full spectra, whereas PCR was applied on full spectra (WinISI software, Infrasoft International, Port Matilda, PA, USA). Prediction equations were validated using a 5-fold cross-validation. Samples with a predicted value that differed more than 2.5 SD from the reference value (T-statistics) were considered outliers and removed from the dataset. Several combinations of scatter corrections (NONE, no correction; SNV\_DT, standard normal variate and detrending; MSC, multiplicative scatter correction) and derivative mathematical treatments (0,0,1,1; 1,4,4,1; 2,5,5,1; where the first digit is the number of the derivative, the second is the gap over which the derivative is calculated, the third is the number of data points in the first smoothing and the fourth is the number of data points in the second smoothing) were tested. The performances of the prediction models were evaluated based on the number of the standard error of calibration (SEC), cross-validation (SECV), the coefficient of determination of cross-validation ( $R^2_{cv}$ ) and the ratio performance to deviation of cross-validation (RPD<sub>cv</sub>) calculated as the ratio between SD and SECV [43]. Predictions were considered excellent when  $R^2$  was greater than 0.91, good when  $R^2$  ranged from 0.82 to 0.90, approximate when  $R^2$  was between 0.66 and 0.81, and poor when  $R^2$  was less than 0.66 [44]. Prediction models with RPD greater than 2.5 were considered adequate for analytical purposes [45], whereas prediction models with RPD smaller than 1.5 were considered unsatisfactory [44].

## 3. Results and Discussion

### 3.1. Chemical Properties

Shell content was on average 13.25% (Table 1), with minimum and maximum values (11.13% and 18.34%, respectively) in line with those reported in the literature (12–20%) [46,47]. Although the shell provides protection to the nib from mold and insects infestations, the shell content should be as low as possible (10–14%) because it has very little commercial value for the cocoa processor: it is removed during cocoa bean processing and it mainly constitutes a waste material [48].

**Table 1.** Descriptive statistics of cocoa beans: SD = standard deviation; CV = coefficient of variation; TPC = total phenolic component; TA = titratable acidity; FI = fermentation index; DM = dry matter.

	Minimum	Maximum	Mean	SD	CV (%)
Shell (%)	11.13	18.34	13.25	1.54	11.59
Fat (%)	36.96	48.39	44.72	1.94	4.35
Protein (%)	8.32	15.43	13.85	1.13	8.14
TPC (mg/g dry defatted powder)	32.58	98.04	56.42	13.32	23.60
pH	4.84	6.47	5.58	0.36	6.50
TA (mMol NaOH/100 g powder)	8.20	26.81	17.19	4.22	24.52
FI (A460/A530)	0.57	2.24	1.29	0.49	38.43
DM (%)	93.30	95.76	94.51	0.59	0.62
Ash (%)	2.34	3.66	2.99	0.30	10.10

Dry matter was on average 94.51%, with a minimum of 93.30%. These values correspond to an average moisture content of 5.49% and maximum of 6.70%, which are mainly below the optimal commercial levels of 6.5–8.0% as reported in CAOBISCO/ECA/FCC [19] but are in line with data found in the literature [49]. Moisture is a parameter that depends on storage conditions: since storage conditions of the studied samples varied, this may have affected the final moisture levels.

The average ash content of 2.99% found in our samples was in line with data reported in the literature [48,50]. With regards to fat content, which is the most abundant macronutrient in cocoa beans, only one sample presented a value below 40 g/100 g (i.e., 36.96 g/100 g), while the average fat content was 44.72 g/100 g. These data are in line with other studies [21,50]. African cocoa beans have generally higher fat content than American beans [16], but this was not observable in our set of samples. However, according to literature, the fat content can vary greatly from values of about 40 g/100 g to values of 57–58 g/100 g depending on different factors such as: genotype, plant age, growing practices, fermentation, drying processes and environmental conditions [51,52].

FI is one of the most used parameters for determining the degree of fermentation of cocoa beans as an indirect measure of the anthocyanin content [29,35]. In our case study, 22 out of 56 samples had a FI slightly below 1, with a minimum value of 0.57, which would indicate a low fermentation degree. The maximum value was 2.24 and the average was 1.29. The coefficient of variation for this parameter was particularly high (38.43%). Since the FI is an indirect measurement of anthocyanin content, the high dispersion of data might be due to factors other than solely the fermentation degree. It has been reported that different hybrids or genotypes have different pigments and that phenolic compounds are quantitatively affected by cocoa growth conditions (microclimate and position of pods on the tree) [29].

The TPC in the dried fat-free mass of our samples exhibited a wide variation, ranging from 32.58 to 98.04 mg/g dry defatted powder. In fermented beans, TPC should be approximately 5% in the dried fat-free mass, and values above 10% are considered a sign of a bad fermentation [53]. The average value of TPC in our samples was 56.42 mg/g dry defatted powder (equals to 5.6%) that would indicate well-fermented beans. Moreover, few samples showed values close to 10%. Overall, the values are in line with those reported in Anyidoho, et al. [54] and Djikeng, et al. [55].

In dried cocoa beans, a high degree of acidity is usually associated with a pH of 5.0 or less [19]. Some studies report that beans of higher pH (5.5–5.8) are considered unfermented, with a low fermentation index, and result in chocolates with high astringency [32], while beans of lower pH (4.75–5.19) are considered as well-fermented. Other studies report that pH of 5–6 is considered good for flavor development, and cocoa beans with pH below 4.5 are not accepted by cocoa bean processors because they show low levels of flavor precursors, and high acidic-derived products [35]. The pH can still be considered as a good indicator of fermentation as higher pH correlates to a lower fermentation degree [16] and an “international acceptable range” of 5.00–5.55 for dried cocoa beans [56] can be

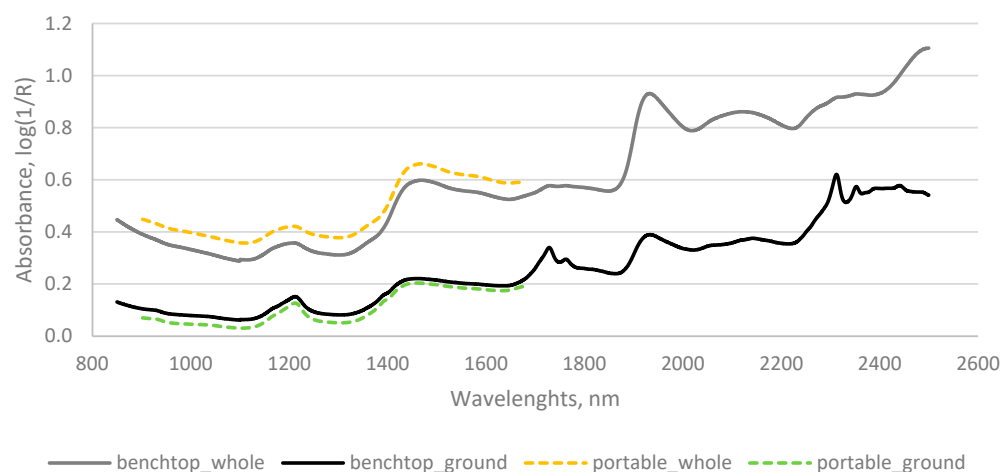
considered as a valid reference. In our case study, cocoa beans had an average pH of 5.58 with a minimum of 4.84. This describes a situation of well fermented samples.

The titratable acidity value is often associated with the beans' pH. The present results confirm an overall good fermentation of the samples with an average titratable acidity of 17.19 mmol NaOH/100 g powder, in line with data reported in the literature [57,58].

Overall, this set of samples included many variation factors (e.g., genetic variety, crop, fermentation and drying conditions, transport, and storage) giving rise to high coefficients of variation in most of the studied parameters [59].

### 3.2. Spectral Characteristics of Cocoa Samples

Figure 1 is representative for average NIR spectra of cocoa beans samples obtained by FOSS DS 2500. The spectra show high similarity with spectra found in the literature [21,27,50,51]. Since cocoa beans contain about 50% of fat (Table 1), absorption spectra are dominated by signals derived from C=O and CH<sub>2</sub> groups [49]. The absorptions around 1930 nm are caused by the second overtone vibration of ester C=O and O–H asymmetric stretching [49,60]. Caporaso, et al. [61] reported that wavelength of 1919 nm has been attributed to the C=O stretching second overtone in the carbonyl groups (–CO<sub>2</sub>H or CONH) but this absorption band is very close to 1923 nm, which is assigned to the O–H group of water and therefore it might be influenced by this group.



**Figure 1.** NIR spectra (mean) of whole (gray) and ground (black) cocoa beans acquired with benchtop spectrometer (FOSS DS 2500) and NIR spectra (mean) of whole (yellow) and ground (green) cocoa beans acquired with portable NIR spectrometer (PoliSPEC-NIR).

The combination vibrations of CH<sub>2</sub> stretch and CH<sub>2</sub> deformation appear around 2320 nm. Moreover, the absorption at 1744 nm has been previously assigned to C–H stretch first overtone (CH<sub>2</sub>) of lipids, and the CH<sub>2</sub> group also absorbs at 1725 nm, due to the C–H stretch first overtone [61]. Similar wavelength values (i.e., 1750 nm and 1730 nm), associated with first overtones of symmetric and anti-symmetric C–H stretch vibration (CH<sub>2</sub>-groups), are reported by Krahmer et al. [49].

Fat content is also related to the absorption bands visible around 1200 nm, as reported by Hayati et al. [27]. The authors also argued that the bands in the wavelength regions of 1460–1490 nm and 1920–1980 nm are most likely related to moisture content (O–H bonds). However, absorbances around 1450 nm have been attributed to carbonyl groups (e.g., ketones and aldehydes) as well as O–H polymeric groups, which can be due to complex carbohydrates, and the region between 1400 nm and 1440 nm has also been attributed to aliphatic alcohols and phenols [61].

Absorbance around 1490 nm has been attributed in the literature to several possible chemical bond vibrations, including N–H stretch first overtone and O–H stretch first overtone, thus indicating amides or compounds such as cellulose [61]. Accordingly,



Krahmer et al. [49] reported that first overtones of intermolecular H-bridges and stretch vibrations of amidic NH-groups can be observed in the region of 1400 to 1500 nm and the corresponding combination of two amides can be found around 2130 nm.

Barbin et al. [50] associated the broad peaks around 1190, 1460 and 1950 nm with O–H, C–H, N–H stretch first and second overtones and combination bands that can be attributed to water absorption and protein changes.

Peaks around 1215 nm are visible and are associated with –CH=CH second overtone [23] and even C–H stretching second overtone (–CH<sub>3</sub> or –CH<sub>2</sub>) of carbohydrates is associated with this wavelength [61].

The absorbance at 2057 nm indicates an N–H stretch/amide 1st combination band, which has been attributed to protein, while the peaks at 2145 and 2313 nm have been tentatively attributed to C–H deformation and C–H deformation and C–H bend second overtones respectively, both indicating lipids [61].

### 3.3. Calibration Models for Cocoa Beans Quality

Variable selection is generally applied in the multivariate analysis to extract the most informative region, removing redundant information. However, among the approaches tested in this study, a lower prediction was observed for the PCR than the MPLS approach as observed in the study of Xie et al. [37]. In detail, in the present study, the PCR showed poor performance of prediction for all traits investigated (see Supplementary Table S2).

Comparing the performance of prediction using the MPLS between full and iPLS selected spectra, it was observed that among the eight parameters, the best prediction was achieved using the full spectra for seven of them (see Supplementary Table S3). The iPLS wavelength selection had a better performance in the fat prediction ( $R^2_{cv}$  of 0.86 and RPD of 2.88) that did not differ substantially from the prediction obtained using the whole spectrum (900–1680 nm;  $R^2_{cv}$  = 0.83 and RPD = 2.43).

The results of prediction performance for the benchtop (NIR FOSS DS 2500) and the portable (PoliSPEC-NIR) spectrometers are presented in Tables 2 and 3, which describe data obtained from whole cocoa beans and peeled-ground cocoa beans, respectively.

**Table 2.** Fitting statistics of prediction models for whole cocoa traits developed using cross-validation results for benchtop (NIR FOSS DS 2500) and portable (PoliSPEC-NIR) NIR-spectrometers.

FOSS DS 2500										
Math Treatment		Constituent	N	Mean	SD	SEcal	R <sup>2</sup> cal	SEcv	R <sup>2</sup> cv	RPD
NONE	1441	Fat	52	2.48	0.26	0.13	0.77	0.15	0.69	1.78
NONE	2551	Protein	51	0.77	0.09	0.04	0.81	0.06	0.66	1.71
SNV_DET	2551	TPC	54	3.04	0.66	0.61	0.15	0.67	0.03	0.98
NONE	0011	pH	54	5.60	0.36	0.21	0.65	0.24	0.58	1.52
MSC	0011	TA	53	16.81	3.99	2.73	0.53	2.98	0.46	1.34
MSC	2551	FI	55	0.07	0.03	0.02	0.25	0.03	0.07	1.03
NONE	1441	DM	55	94.49	0.57	0.26	0.80	0.31	0.72	1.86
MSC	2551	Ash	54	0.16	0.03	0.01	0.87	0.02	0.51	1.43
PoliSPEC-NIR										
Math Treatment		Constituent	N	Mean	SD	SEcal	R <sup>2</sup> cal	SEcv	R <sup>2</sup> cv	RPD
MSC	1441	Fat	54	2.45	0.27	0.18	0.57	0.21	0.38	1.28
SNV_DET	0011	Protein	50	0.78	0.10	0.04	0.83	0.07	0.56	1.49
MSC	2551	TPC	48	3.04	0.60	0.22	0.87	0.40	0.56	1.51
MSC	0011	pH	50	5.56	0.37	0.18	0.76	0.20	0.70	1.83
MSC	1441	TA	55	17.26	4.22	2.77	0.57	3.03	0.48	1.39
NONE	0011	FI	55	0.07	0.03	0.02	0.66	0.02	0.58	1.55
MSC	0011	DM	49	94.50	0.54	0.26	0.77	0.32	0.66	1.72
SNV_DET	0011	Ash	53	0.16	0.03	0.02	0.66	0.02	0.45	1.36

NONE = no correction; SNV\_DET = SNV and detrend; MSC = multiplicative scatter correction; SD = standard deviation of reference data selected; SEcal = standard error in calibration; R<sup>2</sup>cal = coefficient of determination of calibration; SEcv = standard error in cross-validation; R<sup>2</sup>cv = coefficient of determination of cross-validation. TPC = total phenolic compound; TA = titratable acidity; FI = fermentation index; DM = dry matter.

**Table 3.** Fitting statistics of prediction models for ground cocoa traits developed using cross-validation results for benchtop (NIR FOSS DS 2500) and portable (PoliSpec NIR) NIR-spectrometers.

FOSS DS 2500										
Math	Treatment	Constituent	N	Mean	SD	SEcal	R <sup>2</sup> cal	SEcv	R <sup>2</sup> cv	RPD
NONE	1441	Fat	56	2.45	0.27	0.11	0.84	0.13	0.76	2.11
SNV_DET	0011	Protein	54	0.77	0.11	0.02	0.95	0.03	0.91	3.40
NONE	2551	TPC	55	3.04	0.65	0.51	0.40	0.59	0.16	1.10
MSC	0011	pH	50	5.57	0.37	0.08	0.95	0.13	0.88	2.96
NONE	1441	TA	52	16.69	3.86	0.80	0.96	1.43	0.86	2.70
NONE	1441	FI	55	1.29	0.50	0.27	0.70	0.38	0.42	1.31
MSC	0011	DM	56	94.51	0.59	0.15	0.94	0.18	0.90	3.20
MSC	0011	Ash	50	0.16	0.03	0.01	0.90	0.01	0.89	2.98
PoliSPEC-NIR										
Math	Treatment	Constituent	N	Mean	SD	SEcal	R <sup>2</sup> cal	SEcv	R <sup>2</sup> cv	RPD
NONE	2551	Fat	52	2.45	0.27	0.09	0.88	0.12	0.82	2.34
NONE	0011	Protein	53	0.77	0.11	0.04	0.84	0.05	0.79	2.17
MSC	0011	TPC	55	3.04	0.65	0.50	0.42	0.57	0.23	1.14
NONE	1441	pH	53	5.56	0.36	0.11	0.90	0.18	0.74	1.98
NONE	0011	TA	51	17.23	4.07	1.14	0.92	1.76	0.81	2.32
NONE	0011	FI	56	1.29	0.49	0.41	0.33	0.42	0.26	1.17
NONE	0011	DM	54	94.54	0.59	0.25	0.81	0.27	0.79	2.17
NONE	1441	Ash	54	0.16	0.03	0.01	0.89	0.01	0.76	2.08

NONE = no correction; SNV\_DET = SNV and detrend; MSC = multiplicative scatter correction; SD = standard deviation of reference data selected; SEcal = standard error in calibration; R<sup>2</sup>cal = coefficient of determination of calibration; SEcv = standard error in cross-validation; R<sup>2</sup>cv = coefficient of determination of cross-validation. TPC = total phenolic compound; TA = titratable acidity; FI = fermentation index; DM = dry matter.

Generally, most of the cocoa studies were performed on ground cocoa to reduce the effects of the physical sample properties on spectra collection [24]. Indeed, for both NIRS devices, the best performances of prediction were observed on ground sample, probably due to the enhanced homogeneity of the samples characterized by a similarity in the particles size and in a more compacted powder that affects the scattering of light.

In this study, spectra corrections by mathematical treatments to remove irrelevant data such as noise and background information were evaluated. In particular, SNV and MSC were used as pre-processed methods to remove the influence of solid particle size and the surface scattering; moreover, the methods above are mainly recognized as the best mathematical treatment in the equation models developed for whole cocoa. The SNV\_DT and MSC treatments improved the prediction accuracy for some quality parameters of both whole and ground cocoa bean samples, while for other parameters raw spectra gave the best results. This was in line with Barbin et al. [50] who found no considerable improvement of the predictive ability when comparing different pre-processing methods with the original raw data. Indeed, Barbin et al. [50] stated that since the complexity of the models was similar to that obtained with the original data, it is feasible to use the raw spectra to build prediction models for both whole beans and ground cocoa samples.

Moreover, to evaluate the performance of technologies on the market, the whole spectrum was considered to perform the prediction equations, although some researchers suggest that selection of spectral intervals could lead to higher prediction performances [35,62]. All the predictions performed against whole bean sample can be considered as approximate to poor [44] with the highest capability achieved for DM (R<sup>2</sup>cv = 0.72; RPDcv = 1.86) for the benchtop and for pH (R<sup>2</sup>cv = 0.70; RPDcv = 1.83) with portable device (Table 2). In general, the minor prediction capability in whole cocoa beans compared to the ground sample has been confirmed also in the study of Hernández-Hernández et al. [63], in which the poor performance of chemical predictions was attributed to the shell that reflects the incident light hindering the interaction with internal constituents. Although predictions on whole cocoa beans were not adequate for quantitative purposes, they could represent a fast

approach for food business operators to sort cocoa beans towards a specific transformation according to high or low value. Moreover, at germplasm banks and breeding programs, a rapid whole cocoa analysis reduces the time required for the shell removing (usually carried out by hand in the laboratory), suggesting NIRS devices are capable to identify functional genotypes to improve qualitative aspects in cocoa products [63].

Excellent performance was obtained in ground cocoa for protein content ( $R^2_{cv} = 0.91$ ;  $RPD_{cv} = 3.40$ ) and very good prediction was achieved for DM ( $R^2_{cv} = 0.90$ ;  $RPD_{cv} = 3.20$ ), ash ( $R^2_{cv} = 0.89$ ;  $RPD_{cv} = 2.98$ ), pH ( $R^2_{cv} = 0.88$ ;  $RPD_{cv} = 2.96$ ) and TA ( $R^2_{cv} = 0.86$ ;  $RPD_{cv} = 2.70$ ) using the NIR FOSS DS 2500 spectrometer (850–2500 nm) (Table 3). The PoliSPEC-NIR spectrometer (900–1680 nm) had the best predicting performances for fat content ( $R^2_{cv} = 0.82$ ;  $RPD_{cv} = 2.34$ ) in ground samples; whereas, for the other traits, the portable device showed lower performances compared to the benchtop (Table 3).

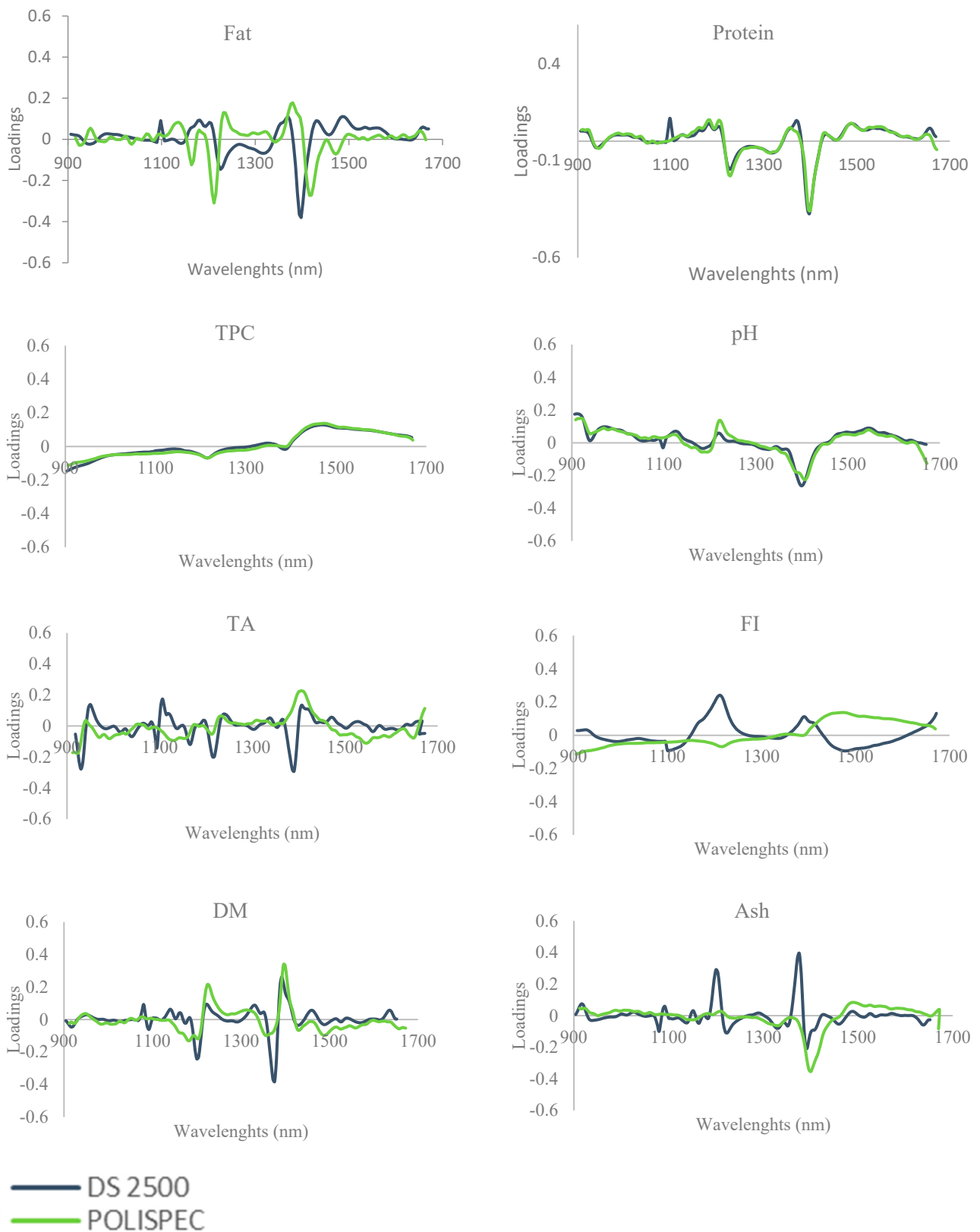
To deeper investigate if the divergences between the devices might depend on the different spectral range used, a further prediction equation was performed for the benchtop using the same spectral range (900–1680 nm, every 2 nm) of the portable tool. In the comparison with the performance obtained considering the whole spectrum, a greater performance of predictions was observed for ash ( $R^2_{cv} = 0.90$ ;  $RPD_{cv} = 3.20$ ), protein ( $R^2_{cv} = 0.93$ ;  $RPD_{cv} = 3.84$ ), DM ( $R^2_{cv} = 0.94$ ;  $RPD_{cv} = 4.16$ ), and lipids ( $R^2_{cv} = 0.83$ ;  $RPD_{cv} = 2.43$ ). However, although the TPC remained unpredictable, an increment was observed in the new prediction equation ( $R^2_{cv} = 0.46$ ;  $RPD_{cv} = 1.37$ ). Although a good predictive capability was maintained, lower performance prediction was observed for TA ( $R^2_{cv} = 0.85$ ;  $RPD_{cv} = 2.60$ ) and pH ( $R^2_{cv} = 0.82$ ;  $RPD_{cv} = 2.34$ ).

Thus, to comprehend the origin of the performance divergences between devices, the component loadings were developed for each tool to assess and compare the interactions between wavelengths and functional groups (Figure 2). The loading plots permit to better understand which wavelengths are more informative for a specific trait variability, showing the range which is mostly considered to develop the model. A strong similarity between portable and benchtop devices were overall observed for chemical parameters directly quantified.

In particular, although the same ranges and performance of prediction ( $R^2_{cv} = 0.83$ ) were obtained in both devices for lipid loading plot, the highest loadings were observed in the spectral region between the 1212 and 1232 nm and 1368 and 1398 nm for portable and benchtop, respectively.

Such association between those range and lipid variability has been confirmed by [64] in cereal food products. Similar patterns for the protein loading plot were observed between the two devices; however, the high loadings observed between 1200 to 1400 nm were related to C–H second overtone and N–H stretching first overtone of protein, respectively [65,66]. Moreover, a high loading was observed around 1100 nm exclusively for the benchtop device; this is probably due to the higher sensitivity of the device that is reflected in the best performance of prediction ( $R^2_{cv} = 0.93$ ) the range 1100–1400 nm being considered as an essential spectral region for the protein quantification analysis [67].

A comparable loading plot was also observed for pH in which the highest trait variability was explained by the 910 [68] and 1398 nm for both devices. Divergences in titratable acidity loading patterns were found; however, the most informative wavelengths (930–950; 1106; 1390–1400 nm) are related to the second combination region of the carboxylic acids [69]. The loading plot of DM showed notable peaks between 1200–1224 and 1373–1394 nm, mainly related to the water [69].



**Figure 2.** Loadings for the first principal component of fat, protein, pH, titratable acidity (TA), total phenolic compound (TPC), fermentation index (FI), dry matter (DM), and ash for NIR spectra of the ground cocoa samples for DS 2500 (blue line) and PoliSPEC NIR (green line).

Ash being an inorganic matter cannot be directly detected by NIRS; its amount is indirectly measured by the association with organic bonds, thus the loadings plot and the highest variability observed for ash is 1200 nm and 1376 nm for the benchtop, and 1396 nm for the portable device account for other organic components. Otherwise, loadings plots observed for TPC and FI were not strictly related to a specific spectrum range, probably due to the lower variability collected with the samples considered. In general, the performance divergences between the two NIRS devices could be explained by the difference in the detector equipment; in detail, the semiconductors included in portable (PoliSPEC-NIR) and benchtop (NIR FOSS DS 2500) devices are Indium gallium arsenide (InGaAs) and silicon lead sulfide, respectively, which affect the spectral response and the prediction capability Lin, et al. [70].

In our study, the accuracy of prediction for both FI and TPC was not satisfactory for any of the instruments and for both whole and ground cocoa bean samples. The influence of variable fermentation degrees of cocoa samples can be crucial in the prediction of FI and TPC, which are strictly related to the fermentation level of cocoa beans. Sunoj, Igathinathane and Visvanathan [32] showed how factors such as pod storage duration (before the fermentation process), and fermentation time, had a significant effect on the fermentation index, which was seen to increase together with the increment of these two parameters. The authors argued that these parameters are indirectly affected by the samples' chemical composition, thus the accuracy of prediction models are generally lower than those reported for major components. The reason might fall on the fact that our samples included only commercial cocoa beans which were supposed to be well-fermented, although with some natural variations, thus reducing the variability for the TPC and FI. Moreover, there could have been a negative influence of lipid absorbances in the models for TPC: fat has been indicated as a disturbance factor as beans with higher relative fat content have lower non-fat solids, where polyphenols are concentrated [61].

Although the FI was not correctly predicted by the constructed models, the estimations in ground samples of parameters related to correct fermentation such as pH and TA were approximative and good with the portable and benchtop devices, respectively, in line with previous results [25,49]. This method could provide a rapid and low-cost multiparametric analysis for cocoa evaluation. Portable instruments are usually less expensive than benchtop solutions (about a fifth) [71], and the cost of analyses are mainly related to the development and upgrade of calibration curves. Moreover, compared to wet analyses, through the application of spectrometric methods the cost of the analytical determination is drastically reduced as the number of examined samples increases.

The presented prediction models might be the basis for an overall cocoa bean quality evaluation based on NIR spectra. However, despite the presented parameters being good indicators of cocoa bean quality, a grading classification of cocoa beans' was beyond the scope of the present work, as it would require the investigation of other indicators, also related to the sensorial profile of the beans, as reported in previous studies on cocoa quality indexes (CQI) [72,73].

#### 4. Conclusions

The results of this paper demonstrated that NIRS portable and benchtop devices coupled with chemometrics methods could be adopted for the chemical evaluation of commercial cocoa beans. The performances of predictions are affected by the presence of shell and the sample particle sizes of cocoa beans. The current study has successfully demonstrated that NIR, as a nondestructive analytical method, can be considered as rapid and reliable option to traditional methods to quantify lipids, protein, pH, titratable acidity, dry matter and ash in cocoa ground beans.

The NIRS benchtop instrument provided better performance of quantification considering the whole (800–2500 nm) and the reduced spectrum (900–1680 nm) than the portable device. Variable selection through iPLS or PCR did not improve prediction models compared to full spectra analyses. Benchtop instrument showed excellent prediction capability



in DM ( $R^2_{cv} = 0.94$ ), protein ( $R^2_{cv} = 0.93$ ) and ash ( $R^2_{cv} = 0.90$ ), whereas lipids ( $R^2_{cv} = 0.83$ ), TA ( $R^2_{cv} = 0.86$ ) and pH ( $R^2_{cv} = 0.88$ ) were well predicted on ground beans considering wavelengths between 900–1680 nm. Those results indicate that models developed for benchtop devices are applicable for cocoa quality control as an excellent option to substitute conventional methods.

On the other hand, the NIRS portable device showed lower but valuable performance of prediction than benchtop spectrometer. The prediction obtained for handheld device represents an appealing strategy for food business operators to apply in the field to control and check the product in every phase of trade and transportation, and also to segregate whole cocoa beans targeted to a specific transformation in different supply chains.

Based on these results, further studies including a wider variability of fermentation phases, cocoa bean varieties and origins as well as additional production steps of the cocoa supply chain could be investigated to support the fair-trade cocoa sector.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/xxx/s1>, Table S1: Cocoa origin of different commercial lot analysed in this study Table S2: Fitting statistics of prediction models for ground cocoa traits developed using full spectra and principal component regression (PCR) and cross-validation results for benchtop (NIR FOSS DS 2500); Table S3. Fitting statistics of prediction models for ground cocoa traits developed using selected wavelengths through the interval PLS (iPLS) and cross-validation results for benchtop (NIR FOSS DS 2500).

**Author Contributions:** Conceptualization, P.C., M.F., K.D. and L.F.; methodology, D.V.d.W., K.D., M.M.; formal analysis, M.F., S.C. and M.M.; resources, P.C. and L.F.; data curation, M.F., S.C. and M.M.; writing—original draft preparation, M.F. and S.C.; writing—review and editing, S.C., D.V.d.W., K.D., L.F. and P.C.; supervision, P.C. and L.F. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** All related data and methods are presented in this paper. Additional inquiries should be addressed to the corresponding author.

**Acknowledgments:** The authors gratefully acknowledge Roberta Marcuzzi for the valuable help in performing the analyses.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. McNeil, C.L. *Chocolate in Mesoamerica: A Cultural History of Cacao*; University Press of Florida: Gainesville, FL, USA, 2009.
2. Afoakwa, E.O. *Chocolate Science and Technology*; John Wiley & Sons: Hoboken, NJ, USA, 2016.
3. Benea-Popușoi, E.; Casapu, M. The forms of international specialization and ensuing economic flows. Case study on the cocoa products. In *Center for Studies in European Integration Working Papers Series*; Center for Studies in European Integration: Chisnau, Moldova, 2021.
4. Sayago-Ayerdi, S.; García-Martínez, D.L.; Ramírez-Castillo, A.C.; Ramírez-Concepción, H.R.; Viuda-Martos, M. Tropical Fruits and Their Co-Products as Bioactive Compounds and Their Health Effects: A Review. *Foods* **2021**, *10*, 1952. [[CrossRef](#)]
5. Badrie, N.; Bekele, F.; Sikora, E.; Sikora, M. Cocoa agronomy, quality, nutritional, and health aspects. *Crit. Rev. Food Sci. Nutr.* **2015**, *55*, 620–659. [[CrossRef](#)] [[PubMed](#)]
6. Gayi, S.K.; Tsowou, K. *Cocoa Industry: Integrating Small Farmers into the Global Value Chain*; United Nations: New York, NY, USA, 2017.
7. Beg, M.S.; Ahmad, S.; Jan, K.; Bashir, K. Status, supply chain and processing of cocoa—A review. *Trends Food Sci. Technol.* **2017**, *66*, 108–116. [[CrossRef](#)]
8. Dzelagha, B.F.; Ngwa, N.M.; Nde Bup, D. A Review of Cocoa Drying Technologies and the Effect on Bean Quality Parameters. *Int. J. Food Sci.* **2020**, *2020*, 8830127. [[CrossRef](#)]
9. Voora, V.; Bermúdez, S.; Larrea, C. *Global Market Report: Cocoa*; International Institute for Sustainable Development: Winnipeg, MB, Canada, 2019.
10. Villacis, A.H.; Alwang, J.R.; Barrera, V.; Dominguez, J. Prices, specialty varieties, and postharvest practices: Insights from cacao value chains in Ecuador. *Agribusiness* **2021**, *38*, 426–458. [[CrossRef](#)]

11. Lima, L.J.; Almeida, M.H.; Nout, M.R.; Zwietering, M.H. *Theobroma cacao* L., “The food of the Gods”: Quality determinants of commercial cocoa beans, with particular reference to the impact of fermentation. *Crit. Rev. Food Sci. Nutr.* **2011**, *51*, 731–761. [[CrossRef](#)] [[PubMed](#)]
12. Hernandez, C.E.; Granados, L. Quality differentiation of cocoa beans: Implications for geographical indications. *J. Sci. Food Agric.* **2021**, *101*, 3993–4002. [[CrossRef](#)]
13. Santander Muñoz, M.; Rodríguez Cortina, J.; Vaillant, F.E.; Escobar Parra, S. An overview of the physical and biochemical transformation of cocoa seeds to beans and to chocolate: Flavor formation. *Crit. Rev. Food Sci. Nutr.* **2020**, *60*, 1593–1613. [[CrossRef](#)]
14. Opoku-Ameyaw, K.; Baah, F.; Gyedu-Akoto, E.; Anchirinah, V.; Dzahini-Obiatey, H.; Cudjoe, A.; Opoku, S. *Cocoa Manual—A Source Book for Sustainable Cocoa Production*; Cocoa Research Institute of Ghana: Tafo, Ghana, 2010.
15. Saltini, R.; Akkerman, R.; Frosch, S. Optimizing chocolate production through traceability: A review of the influence of farming practices on cocoa bean quality. *Food Control.* **2013**, *29*, 167–187. [[CrossRef](#)]
16. Beckett, S.T. *Industrial Chocolate Manufacture and Use*; John Wiley & Sons: Hoboken, NJ, USA, 2011.
17. Jalil, A.M.; Ismail, A. Polyphenols in cocoa and cocoa products: Is there a link between antioxidant properties and health? *Molecules* **2008**, *13*, 2190–2219. [[CrossRef](#)]
18. Kongor, J.E.; Hinneh, M.; Van de Walle, D.; Afoakwa, E.O.; Boeckx, P.; Dewettinck, K. Factors influencing quality variation in cocoa (*Theobroma cacao*) bean flavour profile—A review. *Food Res. Int.* **2016**, *82*, 44–52. [[CrossRef](#)]
19. CAOBISCO; ECA; FCC. *Cocoa Beans: Chocolate and Cocoa Industry Quality Requirements*; End, M.J., Dand, R., Eds.; ECA-Caobisco-FCC Cocoa Research Fund: Brussels, Belgium, 2015.
20. Scheu, J.; Dand, R. *Cocoa. A Guide to Trade Practices*; International Trade Centre UNCTAD/WTO: Geneva, Switzerland, 2001.
21. Hashimoto, J.C.; Lima, J.C.; Celeghini, R.M.; Nogueira, A.B.; Efraim, P.; Poppi, R.J.; Pallone, J.A. Quality control of commercial cocoa beans (*Theobroma cacao* L.) by near-infrared spectroscopy. *Food Anal. Methods* **2018**, *11*, 1510–1517. [[CrossRef](#)]
22. Fountain, A.C.; Huetz-Adams, F. *Cocoa Barometer 2020*; VOICE Network: Ede, The Netherlands, 2020.
23. Teye, E.; Anyidoho, E.; Agbemafle, R.; Sam-Amoah, L.K.; Elliott, C. Cocoa bean and cocoa bean products quality evaluation by NIR spectroscopy and chemometrics: A review. *Infrared Phys. Technol.* **2020**, *104*, 103127. [[CrossRef](#)]
24. Quelal-Vasconez, M.A.; Lerma-Garcia, M.J.; Perez-Esteve, E.; Talens, P.; Barat, J.M. Roadmap of cocoa quality and authenticity control in the industry: A review of conventional and alternative methods. *Compr. Rev. Food Sci. Food Saf.* **2020**, *19*, 448–478. [[CrossRef](#)]
25. Anyidoho, E.K.; Teye, E.; Agbemafle, R. Differentiation of Organic Cocoa Beans and Conventional Ones by Using Handheld NIR Spectroscopy and Multivariate Classification Techniques. *Int. J. Food Sci.* **2021**, *2021*, 1844675. [[CrossRef](#)]
26. Anyidoho, E.K.; Teye, E.; Agbemafle, R.; Amuah, C.L.Y.; Boadu, V.G. Application of portable near infrared spectroscopy for classifying and quantifying cocoa bean quality parameters. *J. Food Process. Preserv.* **2021**, *45*, e15445. [[CrossRef](#)]
27. Hayati, R.; Zulfahrizal, Z.; Munawar, A.A. Robust prediction performance of inner quality attributes in intact cocoa beans using near infrared spectroscopy and multivariate analysis. *Heliyon* **2021**, *7*, e06286. [[CrossRef](#)]
28. Singleton, V.L.; Rossi, J.A. Colorimetry of total phenolics with phosphomolybdic-phosphotungstic acid reagents. *Am. J. Enol. Vitic.* **1965**, *16*, 144–158.
29. Romero-Cortes, T.; Salgado-Cervantes, M.A.; Garcia-Alamilla, P.; Garcia-Alvarado, M.A.; Rodriguez-Jimenes Gdel, C.; Hidalgo-Morales, M.; Robles-Olvera, V. Relationship between fermentation index and other biochemical changes evaluated during the fermentation of Mexican cocoa (*Theobroma cacao*) beans. *J. Sci. Food Agric.* **2013**, *93*, 2596–2604. [[CrossRef](#)]
30. Racine, K.C.; Lee, A.H.; Wiersema, B.D.; Huang, H.; Lambert, J.D.; Stewart, A.C.; Neilson, A.P. Development and Characterization of a Pilot-Scale Model Cocoa Fermentation System Suitable for Studying the Impact of Fermentation on Putative Bioactive Compounds and Bioactivity of Cocoa. *Foods* **2019**, *8*, 102. [[CrossRef](#)]
31. Afoakwa, E.O.; Jennifer, Q.; Agnes, S.B.; Jemmy, S.T.; Saalia, F. Influence of pulp-preconditioning and fermentation on fermentative quality and appearance of ghanaiian cocoa (*theobroma cacao*) beans. *Int. Food Res. J.* **2012**, *19*, 127–133.
32. Sunoj, S.; Igathinathane, C.; Visvanathan, R. Nondestructive determination of cocoa bean quality using FT-NIR spectroscopy. *Comput. Electron. Agric.* **2016**, *124*, 234–242. [[CrossRef](#)]
33. Hue, C.; Gunata, Z.; Bergounhou, A.; Assemat, S.; Boulanger, R.; Sauvage, F.X.; Davrieux, F. Near infrared spectroscopy as a new tool to determine cocoa fermentation levels through ammonia nitrogen quantification. *Food Chem.* **2014**, *148*, 240–245. [[CrossRef](#)] [[PubMed](#)]
34. Nazaruddin, R.; Seng, L.K.; Hassan, O.; Said, M. Effect of pulp preconditioning on the content of polyphenols in cocoa beans (*Theobroma Cacao*) during fermentation. *Ind. Crops Prod.* **2006**, *24*, 87–94. [[CrossRef](#)]
35. Teye, E.; Huang, X.; Sam-Amoah, L.K.; Takrama, J.; Boison, D.; Botchway, F.; Kumi, F. Estimating cocoa bean parameters by FT-NIRS and chemometrics analysis. *Food Chem.* **2015**, *176*, 403–410. [[CrossRef](#)]
36. Andersen, C.M.; Bro, R. Variable selection in regression—A Tutorial. *J. Chemom.* **2010**, *24*, 728–737. [[CrossRef](#)]
37. Xie, L.H.; Tang, S.Q.; Chen, N.; Luo, J.; Jiao, G.A.; Shao, G.N.; Wei, X.J.; Hu, P.S. Optimisation of near-infrared reflectance model in measuring protein and amylose content of rice flour. *Food Chem.* **2014**, *142*, 92–100. [[CrossRef](#)]
38. Williams, P. Interpretation of Calibration Evaluation II. The Spectral Basics of Interpretation. *NIR News* **2013**, *24*, 21–25. [[CrossRef](#)]

39. Chadalavada, K.; Anbazhagan, K.; Ndour, A.; Choudhary, S.; Palmer, W.; Flynn, J.R.; Mallayee, S.; Pothu, S.; Prasad, K.; Varijakshapanikar, P.; et al. NIR Instruments and Prediction Methods for Rapid Access to Grain Protein Content in Multiple Cereals. *Sensors* **2022**, *22*, 3710. [[CrossRef](#)]
40. Manuelian, C.L.; Currò, S.; Penasa, M.; Cassandro, M.; De Marchi, M. Prediction of minerals, fatty acid composition and cholesterol content of commercial cheeses by near infrared transmittance spectroscopy. *Int. Dairy J.* **2017**, *71*, 107–113. [[CrossRef](#)]
41. Albanell, E.; Miñarro, B.; Carrasco, N. Detection of low-level gluten content in flour and batter by near infrared reflectance spectroscopy (NIRS). *J. Cereal Sci.* **2012**, *56*, 490–495. [[CrossRef](#)]
42. Moghimi, A.; Aghkhani, M.H.; Sazgarnia, A.; Sarmad, M. Vis/NIR spectroscopy and chemometrics for the prediction of soluble solids content and acidity (pH) of kiwifruit. *Biosyst. Eng.* **2010**, *106*, 295–302. [[CrossRef](#)]
43. Williams, P.C.; Sobering, D.C. Comparison of Commercial near Infrared Transmittance and Reflectance Instruments for Analysis of Whole Grains and Seeds. *J. Near Infrared Spectrosc.* **1993**, *1*, 25–32. [[CrossRef](#)]
44. Karoui, R.; Mouazen, A.M.; Dufour, É.; Schoonheydt, R.; De Baerdemaeker, J. A comparison and joint use of VIS-NIR and MIR spectroscopic methods for the determination of some chemical parameters in soft cheeses at external and central zones: A preliminary study. *Eur. Food Res. Technol.* **2006**, *223*, 363–371. [[CrossRef](#)]
45. Sinnaeve, G.; Dardenne, P.; Agneessens, R.; Biston, R. The Use of near Infrared Spectroscopy for the Analysis of Fresh Grass Silage. *J. Near Infrared Spectrosc.* **1994**, *2*, 79–84. [[CrossRef](#)]
46. Quelal-Vásconez, M.A.; Lerma-García, M.J.; Pérez-Esteve, É.; Arnau-Bonachera, A.; Barat, J.M.; Talens, P. Fast detection of cocoa shell in cocoa powders by near infrared spectroscopy and multivariate analysis. *Food Control.* **2019**, *99*, 68–72. [[CrossRef](#)]
47. Okiyama, D.C.; Navarro, S.L.; Rodrigues, C.E. Cocoa shell and its compounds: Applications in the food industry. *Trends Food Sci. Technol.* **2017**, *63*, 103–112. [[CrossRef](#)]
48. Afoakwa, E.O.; Quao, J.; Takrama, J.; Budu, A.S.; Saalia, F.K. Chemical composition and physical quality characteristics of Ghanaian cocoa beans as affected by pulp pre-conditioning and fermentation. *J. Food Sci. Technol.* **2013**, *50*, 1097–1105. [[CrossRef](#)]
49. Krahmer, A.; Engel, A.; Kadow, D.; Ali, N.; Umaharan, P.; Kroh, L.W.; Schulz, H. Fast and neat-determination of biochemical quality parameters in cocoa using near infrared spectroscopy. *Food Chem.* **2015**, *181*, 152–159. [[CrossRef](#)]
50. Barbin, D.F.; Maciel, L.F.; Bazoni, C.H.V.; da Silva Ribeiro, M.; Carvalho, R.D.S.; da Silva Bispo, E.; Miranda, M.d.P.S.; Hirooka, E.Y. Classification and compositional characterization of different varieties of cocoa beans by near infrared spectroscopy and multivariate statistical analyses. *J. Food Sci. Technol.* **2018**, *55*, 2457–2466. [[CrossRef](#)]
51. Agus, B.A.P.; Mohamad, N.N.; Hussain, N. Composition of unfermented, unroasted, roasted cocoa beans and cocoa shells from Peninsular Malaysia. *J. Food Meas. Charact.* **2018**, *12*, 2581–2589. [[CrossRef](#)]
52. Samaniego, I.; Espín, S.; Quiroz, J.; Rosales, C.; Carrillo, W.; Mena, P.; García-Viguera, C. Effect of the growing area on the fat content and the fatty acid composition of Ecuadorian cocoa beans. *Int. J. Food Sci. Nutr.* **2021**, *72*, 901–911. [[CrossRef](#)] [[PubMed](#)]
53. Wollgast, J.; Anklam, E. Review on polyphenols in *Theobroma cacao*: Changes in composition during the manufacture of chocolate and methodology for identification and quantification. *Food Res. Int.* **2000**, *33*, 423–447. [[CrossRef](#)]
54. Anyidoho, E.K.; Teye, E.; Agbemafle, R. Nondestructive authentication of the regional and geographical origin of cocoa beans by using a handheld NIR spectrometer and multivariate algorithm. *Anal. Methods* **2020**, *12*, 4150–4158. [[CrossRef](#)]
55. Djikeng, F.T.; Teyomnou, W.T.; Tenyang, N.; Tiencheu, B.; Morfor, A.T.; Touko, B.A.H.; Houketchang, S.N.; Boungo, G.T.; Karuna, M.S.L.; Ngoufack, F.Z.; et al. Effect of traditional and oven roasting on the physicochemical properties of fermented cocoa beans. *Heliyon* **2018**, *4*, e00533. [[CrossRef](#)]
56. Ndukwu, M.; Udofia, M. Kinetics of change in colour and some bio-chemical composition during fermentation of cocoa bean. *Cogent Food Agric.* **2016**, *2*, 1268743. [[CrossRef](#)]
57. Hii, C.L.; Abdul Rahman, R.; Jinap, S.; Che Man, Y.B. Quality of cocoa beans dried using a direct solar dryer at different loadings. *J. Sci. Food Agric.* **2006**, *86*, 1237–1243. [[CrossRef](#)]
58. Huang, X.; Teye, E.; Owusu-Sekyere, J.D.; Takrama, J.; Sam-Amoah, L.K.; Yao, L.; Firempong, C.K. Simultaneous measurement of titratable acidity and fermentation index in cocoa beans by electronic tongue together with linear and non-linear multivariate technique. *Food Anal. Methods* **2014**, *7*, 2137–2144. [[CrossRef](#)]
59. Kutsanedzie, F.Y.H.; Chen, Q.; Hassan, M.M.; Yang, M.; Sun, H.; Rahman, M.H. Near infrared system coupled chemometric algorithms for enumeration of total fungi count in cocoa beans neat solution. *Food Chem.* **2018**, *240*, 231–238. [[CrossRef](#)]
60. Teye, E.; Huang, X. Novel prediction of total fat content in cocoa beans by FT-NIR spectroscopy based on effective spectral selection multivariate regression. *Food Anal. Methods* **2015**, *8*, 945–953. [[CrossRef](#)]
61. Caporaso, N.; Whitworth, M.B.; Fowler, M.S.; Fisk, I.D. Hyperspectral imaging for non-destructive prediction of fermentation index, polyphenol content and antioxidant activity in single cocoa beans. *Food Chem.* **2018**, *258*, 343–351. [[CrossRef](#)]
62. Satriyo, P.; Munawar, A.A. Data analysis on near infrared spectroscopy as a part of technology adoption for cocoa farmer in Aceh Province, Indonesia. *Data Brief* **2020**, *29*, 105251.
63. Hernández-Hernández, C.; Fernández-Cabanás, V.M.; Rodríguez-Gutiérrez, G.; Fernández-Prior, Á.; Morales-Sillero, A. Rapid screening of unground cocoa beans based on their content of bioactive compounds by NIR spectroscopy. *Food Control.* **2022**, *131*, 108347. [[CrossRef](#)]
64. Vines, L.L.; Kays, S.E.; Koehler, P.E. Near-Infrared Reflectance Model for the Rapid Prediction of Total Fat in Cereal Foods. *J. Agric. Food Chem.* **2005**, *53*, 1550–1555. [[CrossRef](#)] [[PubMed](#)]

65. Saha, U.; Endale, D.; Tillman, P.G.; Johnson, W.C.; Gaskin, J.; Sonon, L.; Schomberg, H.; Yang, Y. Analysis of Various Quality Attributes of Sunflower and Soybean Plants by Near Infrared Reflectance Spectroscopy: Development and Validation Calibration Models. *Am. J. Anal. Chem.* **2017**, *8*, 462–492. [[CrossRef](#)]
66. Williams, P.; Manley, M.; Antoniszyn, J. *Near Infrared Technology: Getting the Best out of Light*; African Sun Media: Stellenbosch, South Africa, 2019.
67. Delwiche, S.R. Protein Content of Single Kernels of Wheat by Near-Infrared Reflectance Spectroscopy. *J. Cereal Sci.* **1998**, *27*, 241–254. [[CrossRef](#)]
68. Shao, Y.; He, Y. Nondestructive measurement of the internal quality of bayberry juice using Vis/NIR spectroscopy. *J. Food Eng.* **2007**, *79*, 1015–1019. [[CrossRef](#)]
69. Workman, J.J.; Weyer, L. *Practical Guide and Spectral Atlas for Interpretive Near-Infrared Spectroscopy*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2012. [[CrossRef](#)]
70. Lin, M.; A Rasco, B.; G Cavinato, A.; Al-Holy, M. Chapter 6—Infrared (IR) Spectroscopy—Near-Infrared Spectroscopy and Mid-Infrared Spectroscopy. In *Infrared Spectroscopy for Food Quality Analysis and Control*; Sun, D.-W., Ed.; Academic Press: San Diego, CA, USA, 2009; pp. 119–143. [[CrossRef](#)]
71. Sandak, J.; Sandak, A.; Zitek, A.; Hintestoisser, B.; Picchi, G. Development of Low-Cost Portable Spectrometers for Detection of Wood Defects. *Sensors* **2020**, *20*, 545. [[CrossRef](#)]
72. Gil, M.; Jaramillo, Y.; Bedoya, C.; Llano, S.M.; Gallego, V.; Quijano, J.; Londono-Londono, J. Chemometric approaches for postharvest quality tracing of cocoa: An efficient method to distinguish plant material origin. *Heliyon* **2019**, *5*, e01650. [[CrossRef](#)]
73. Araujo, Q.R.d.; Loureiro, G.A.H.d.A.; Baligar, V.C.; Ahnert, D.; Faria, J.C.; Valle, R.R. Cacao quality index for cacao agroecosystems in Bahia, Brazil. *Int. J. Food Prop.* **2019**, *22*, 1799–1814. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.