Safety and quality issues in summer squashes using handheld portable NIRS sensors for real-time decision making and on-vine monitoring

José-Antonio Entrenas[a], Dolores Pérez-Marín[b,*], Irina Torres[a], Ana Garrido-Varo[b], María-Teresa Sánchez[a,*]

[a] *Department of Bromatology and Food Technology, University of Cordoba, Campus of Rabanales, 14071 Córdoba, Spain.*

[b] *Department of Animal Production, University of Cordoba, Campus of Rabanales, 14071 Córdoba, Spain.*

**Corresponding authors*. Tel.: +34 957 212576; fax: 34 957 212000

E-mail addresses: teresa.sanchez@uco.es (M.T. Sánchez) or dcperez@uco.es (D. Pérez-Marín).

Abstract

BACKGROUND: Portable handheld near infrared spectroscopy (NIRS) instruments currently present enormous advantages in terms of size, weight and robustness. They also provide fast, precise information that can be obtained *in situ*, and represent a viable option for controlling vegetable safety and quality during the growth period. The aim of this research was to evaluate three handheld portable NIRS instruments for *in situ* and real time analysis of intact summer squashes. 221 summer squashes were analyzed by traditional methods and used to develop calibration models for morphological, safety and quality parameters. Additionally, the longitudinal distribution of nitrate content in summer squashes weighing over 400 g was also studied, and the evolution of this parameter during the harvest period was also tracked to determine which summer squashes and which zones (peduncle, equatorial or stylar) of the vegetable could be earmarked for baby food production.

RESULTS: The robustness of the calibration models obtained confirmed the expectations raised by NIRS technology for morphological, safety and quality control of individual summer squashes, and the models developed with the MicroNIR-1700 instrument were those which proved more accuracy and precision, being the peduncle zone the part that presents a higher content in nitrates.

CONCLUSIONS: It is in the peduncle zone, therefore, where measurements of this parameter must be carried out to decide on the destination of the harvested product. Additionally, summer squashes picked at the end of the harvest are those which must be used for baby food production.

Keywords: Summer squash; Portable NIR sensor; *In situ* determination; Safety and quality parameters; Nitrate content; Baby foods.

INTRODUCTION

Near infrared spectroscopy, which can be defined as a non-invasive environmentally-friendly technique which combines versatility, speed, ease of use and accurate measurements with the low cost of each analysis, offers the possibility of measuring safety and quality attributes in fruits and vegetables.[1] This facilitates its incorporation at different decision-making steps in the agri-food chain, both in the pre-harvest and harvest periods in the field, and in the post-harvest period, in the processing industry.[1-4]

Currently, portable, handheld and compact-design NIRS instruments are in full development and expansion.[5,6] These portable devices run on batteries and offer huge advantages in terms of size, weight and robustness of the analysis in uncontrolled environmental conditions, ~~since they lack mobile elements,~~ in addition to being cheaper to acquire in comparison with the classic laboratory instruments.

Nowadays, there is a wide range of portable instruments of different types in terms of working spectral range, cost and optical design, which are based on different technologies such as micro-electro-mechanical system (MEMS) or linear variable filters (LVF). They represent a clear evolution in the use of NIRS technology: previously, the sample had to be taken to the lab, but now, an *in situ* analysis is carried out.[7] ~~where the sample is located.~~ Faced with such a wide diversity in the characteristics and features of these portable NIRS sensors, there is a need for them to be evaluated in order to choose which is the most suitable for a certain application or a specific product.

The use of portable NIRS sensors can favor the decision-making process in the horticultural sector, allowing to set the optimum harvest time and carry out harvesting strategies in stages, depending on the industrial destination of the product.[8-10] In particular, in vegetables such as summer squashes, where the nitrate content is a key factor when establishing the destination of the harvested product, the use of a handheld NIRS sensor, *in* *situ*, directly on the plant, would facilitate the selective harvesting of this vegetable for its possible use in making baby foods (the maximum level for nitrates in processed cereal-based foods and baby foods for infants and young children is set at 200 mg NO_3 kg^{-1}), according to the European Union legislation.[11]

Likewise, despite the fact that numerous studies carried out on nitrate accumulation in plants have found that the concentration of this substance depends on a number of different factors – plant biology, daylight intensity, soil type, temperature, humidity, sowing density, plant maturity, vegetation period, harvesting period and nitrogen source[12] –the order of nitrate content accumulation in the different organs has only been established (petiole > leaf > stem > roots > inflorescence > tuber > fruit > seed), while the nature of nitrate accumulation inside the fruit and which edible part contains the nitrate has not been studied.[13] This differs for this parameter and for summer squashes from other quality parameters such as dry matter and soluble solid content (SSC) and from other fruits such as melon, where the variation inside the fruit has been widely researched.[14,15] This is a vital factor in determining the key zones for analysis and also for saving certain parts of this vegetable for more critical destinations, such as baby foods.

Sánchez *et al.*[9] determined safety and quality parameters in summer squashes on the plant, using an NIRS instrument, Phazir 2400, based on MEMS technology. However, the recent arrival of modern commercial sensors has led to the phasing out of many of the earlier models, such as the Phazir 2400 mentioned above. For this reason, the efficiency of these new portable devices in horticultural applications needs to be assessed.

The main objective of this research work is therefore to evaluate and compare handheld, portable NIRS instruments when used to assess the safety and quality of summer squashes on the plant. It also aims to study the longitudinal nitrate accumulation in this fruit to stablish which zone of the vegetable (peduncle, equatorial or stylar) contains a greater accumulation of nitrates and is, therefore, the key zone to be analyzed, and the one which determines the destination of the vegetable in the processing industry.

MATERIALS AND METHODS

Sampling and reference methods

A total of 221 summer squashes (126 below 400 g in weight and 95 over 400 g) *(Cucurbita pepo* subsp. *pepo* morphotype zucchini cv. Mirza), grown in an open-air plantation in the district of La Montiela, Santaella (Córdoba, Spain), were harvested between May and July 2017. On arrival at the laboratory, the fruits were promptly placed in refrigerated storage at 5 ºC and 85% relative humidity. Prior to measurement, each sample was left at room temperature to stabilize at the laboratory temperature of 20 ºC.

The summer squashes were individually weighed on an electronic balance (0–1,000 ± 0.01 g; model P1000 N, Metter-Toledo, GmbH, Greifensee, Switzerland). Their length was measured using a measuring tape and the equatorial diameter was then measured using a digital precision caliper (0–300 ± 0.01 mm; Comecta, Barcelona, Spain).

Nitrate content, dry matter and SSC were measured following Sánchez *et al.*[9]. To analyze these parameters in summer squashes weighing over 400 g, the fruit was divided into three zones: the peduncle zone (upper third of the squash starting at the peduncle), the equatorial zone (middle third in the equator of the fruit) and the stylar zone (lower third of the fruit, starting at the pistil scar). All the analytical measurements were performed immediately after NIR spectrum collection and in duplicate.

Spectral data collection

The NIR spectra of the intact summer squashes were collected in reflectance mode (log 1/R) using three handheld NIRS instruments:

- Phazir 2400, a handheld MEMS-based NIR digital transform spectrometer (Polychromix, Inc., Wilmington, MA, USA). This compact, robust spectrometer weighing 1.7 kg is specially designed for *in situ* NIRS analysis. The equipment scans at a non-constant interval of approximately 8 nm, across the NIR wavelength range of 1600 to 2400 nm, with a window area of around 55 mm^2. The sensor integration time was 600 ms and each spectrum was the mean of 5 scans. This instrument is equipped with special quartz protection to prevent dirt from accumulating. The instrument's performance was checked every 10 min, following the diagnostic protocols provided by the manufacturer, and white reference measurement was carried out using Spectralon™ (a NIR reflectance standard with a 99% diffuse reflectance) as reference. Although the model has now been phased out, it was used as a reference sensor to compare with the other instruments.

- MicroPhazir, a handheld MEMS-based NIR digital transform spectrometer (Polychromix Inc., Wilmington, MA, USA). This model is an updated version of the Phazir 2400 and its instrumental design and optical features are therefore, very similar: it is a pistol-shaped device which is portable, compact and robust. The window area is at around 41 mm^2 and it works in the spectral range of 1600 to 2400 nm with a non-constant interval of 8 nm. However, it is much lighter (1.2 kg) than its predecessor, which makes it more comfortable when analyzing the product. Unlike the former, it has an internal reference which enables easy calibration in the field. The sensor integration time was 600 ms and each spectrum was the mean of 5 scans. The device is equipped with quartz protection to prevent dirt accumulation.

- A MicroNIR-1700 LVF spectrometer (VIAVI Solutions, Inc., San Jose, California, USA). This portable miniature spectrometer is extremely light (64 g, without including the handle of 150 g and the acquisition and data processing device). Its optical window is larger than that of the previous equipments (the measurement area is around 227 mm^2). This microspectrometer covers a 910 to 1676 nm spectral range, with a constant interval of 6.2 nm. The instrument's performance was checked every 10 min. A white reference measurement was obtained using Spectralon™, while a dark reference was obtained from a fixed point in the room. The sensor integration time was 11 ms and each spectrum was the mean of 200 scans.

The main features of these instruments are summarized in Table 1.

To collect NIR spectra using these three spectrometers, the fruits, regardless of weight, were divided into the three zones (peduncle, equatorial and stylar) mentioned above.

Four spectral measurements were taken in each of the three zones analyzed, the first at a random location in the center of the analyzed zone, which were then rotated 90º after each measurement, thus obtaining 12 spectra per summer squash.

The 12 spectra were averaged to provide a mean spectrum per fruit in the case of summer squashes weighing below 400 g (126 spectra), for all the parameters analyzed.

For summer squashes weighing over 400 g, the same procedure as described above for taking the spectra was carried out. To develop predictive models of the morphological (weight, length and equatorial diameter) parameters, an average was taken of the 12 spectra obtained initially, resulting in a single spectrum per fruit, which produced a total of 95 spectra. However, taking into account the fact that the analysis of nitrate content, dry matter and SSC were carried out by zones, the 4 spectra corresponding to each of the studied zones were averaged, thus obtaining an average spectrum per zone – i.e. a total of 285 spectra (95 fruits · 3 zones/fruit · 1 spectrum/zone).

Data processing

Data pre-processing and chemometric treatments were performed using the WinISI II software package version 1.50 (Infrasoft International LLC, Port Matilda, PA, USA).[16] Before the spectral data were processed, a study was conducted to select the most suitable spectral range for the instruments tested to carry out the morphological, safety and quality control of summer squashes. To achieve this, the 1,1,1,1 derivation treatment was applied (the first digit being the number of the derivative, the second the gap over which the derivative is calculated, the third the number of data points in a running average or smoothing, and the fourth the second smoothing) without scatter correction, which highlights the areas of the spectrum where the signal/noise ratio is degraded.

Spectral repeatability

Spectrum quality was evaluated using the root mean square (RMS) statistic. The RMS statistic is defined as the averaged root mean square of differences between the different subsamples scanned at *n* wavelengths.[18,19] This statistic indicates the similarity between different spectra of a single sample. To establish a threshold for this statistic, two strategies were tested. In the first, the repeatability was measured considering only the spectral information collected in the center of the peduncle zone of the summer squashes, after rotating the product 90º between each measurement, using 10 samples for each type of summer squash analyzed. In the second strategy, 12 spectra (3 zones * 4 spectra/zone) were taken following the same procedure and number of samples above mentioned. An admissible limit for spectrum quality and repeatability was set following the procedure described by Martínez *et al.*[20] to calculate the standard deviation limit (STD_{limit}) from the RMS statistic and obtain an RMS cut-off value.

Definition of calibration and validation sets

Prior to carrying out NIRS calibrations, the CENTER algorithm was applied to ensure a structured population selection based solely on spectral information, for the establishment of calibration and validation sets.[21] This algorithm performs an initial principal component analysis to calculate the center of the population and the distance of samples (spectra) from that center in an n dimensional space, using the Mahalanobis distance (GH); samples with a GH value > 4 were considered outliers or anomalous spectra. A combination of mathematical pretreatments, standard normal variate (SNV) and de-trending (DT) was applied for scatter correction,[22] together with the 1,5,5,1 derivate mathematical treatment.[16,18]

To predict the morphological parameters, the CENTER algorithm was applied to the 221 spectra obtained after averaging the 12 spectra taken of each fruit while for the prediction of the safety and quality parameters, and since the analysis was performed by zones in summer squashes weighing over 400 g, the CENTER algorithm was applied to the 411 available spectra.

Having ordered the sample set by spectral distances and once the spectral outliers were removed, NIRS calibration models for the prediction of morphological, safety and quality parameters in intact summer squashes were initially constructed using the calibration sets comprising all the available samples (C1 = 217 samples for morphological parameters and C2 = 407 samples for safety and quality parameters). After analyzing the accuracy and precision of the models obtained and evaluating the three instruments, new calibration models were developed for these parameters using the most suitable instrument. For this purpose, the samples forming the validation set were selected by taking one sample out of every four from the initial sets (C1 and C2). After this procedure, the calibration (C3 = morphological parameters and C4 = nitrate content, dry matter and

SSC) and validation (V3 = morphological parameters and V4 = nitrate content, dry matter and SSC) sets thus comprised the samples shown in Table 2.

Data pre-processing and calibration model construction using a linear regression strategy

NIRS calibration models for the prediction of morphological, safety and quality parameters in intact summer squashes were initially constructed with the calibration sets C1 and C2 respectively, using modified partial least squares regression,[23] with subsequent cross-validation. The calibration set was divided into 4 groups; each group was then validated using a calibration developed for the other samples; finally, validation errors were combined to obtain a standard error of cross-validation (SECV).

For each analytical parameter, different mathematical pretreatments were evaluated. For scatter correction, SNV and DT methods were tested.[22] Additionally, a total of two mathematical derivation treatments were tested: 1,5,5,1; 2,5,5,1.[16,18]

The statistics used to select the best equations were the coefficient of determination for cross-validation (r^2_{cv}), and the ~~standard error of cross-validation~~ SECV. Furthermore, the residual predictive deviation (RPD_{cv}) for cross-validation was calculated as the ratio of the standard deviation (SD) of the reference data to the SECV. This statistic enables SECV to be standardized, facilitating the comparison of results obtained with sets of different means.[24]

Once the best predictive model for each parameter was selected without the elimination of physical-chemical outliers, tests were run for significant differences between models, with a view to identifying the most suitable spectrometer for routine use in on-vine summer squashes during the growing period. The SECV values for the best equations obtained for each parameter with the three instruments were compared using Fisher's F test.[25,26] The values for F were calculated as:

$$F = \frac{(SECV_2)^2}{(SECV_1)^2}$$

where $SECV_1$ and $SECV_2$ are the standard error of cross validation of two different models and $SECV_1 < SECV_2$. F is compared to $F_{critical}$ (1-*P*, n_2-1, n_1-1) read from the table with *P* = 0.05 and n-1 degrees of freedom. If F is higher than $F_{critical}$, the two SECV values are significantly different. When several SECV values are compared, as in this research, a $SECV_{confidence\ limit}$ is calculated using the following formula: $SECV_{confidence\ limit} = SECV_{min} \sqrt{F_{critical}}$ where $SECV_{min}$ is the smallest SECV. As a consequence, none of the models which have a SECV between $SECV_{min}$ and $SECV_{confidence\ limit}$ are significantly different.

Finally, once the best NIRS instrument ~~from the three tested~~ was chosen, new models were developed (optimizing the performance models parameters) with that spectrometer using the C3 and C4 calibration sets. The best-fitting equations obtained for these new calibration sets, as selected by the same statistical criteria mentioned above, were subsequently subjected to external validation using the prediction sets V3 and V4, respectively, following the protocol outlined by Windham *et al.*[27]

Statistical analysis

In order to study the influence of both the harvest date and the zone analyzed, as well as the harvest date x zone interaction, in the nitrate content (wet analysis) of summer squashes weighing over 400 g, a two-factor analysis of variance (ANOVA) was carried out using Statgraphics Centurion XV (StatPoint Inc., Warrenton, North Virginia, USA).

Next, the differences between the means were compared with the Fisher's Least Significant Difference (LSD) test, and differences at $P < 0.05$ were considered to be significant.

RESULTS AND DISCUSSION

Nitrate content evolution during the harvest period

Nitrate content in summer squashes weighing over 400 g were significantly influenced ($P < 0.05$) by harvest date and analyzed zone but not by the harvest date x zone interaction (Table 3). For each of the analyzed zones, the nitrate content decreased significantly ($P < 0.05$) as the harvest period progressed, reaching minimum values on the last harvesting day (07/13/2017). As regards the zone analyzed, the nitrate content was significantly higher ($P < 0.05$) in the peduncle zone, which indicates that it is here where this substance accumulates the most. It is also worth noting that there is a significantly lower nitrate content in the equatorial zone ($P < 0.05$). Therefore, when determining the destination for the summer squashes after harvesting, both farmers and the processing industry should carry out the NIRS and wet analysis to measure the nitrate content present in the peduncle zone of the vegetable. It is also recommended to use end-of-harvest summer squashes to elaborate baby foods (nitrate content < 200 mg kg^{-1}), since the nitrate content in the three analyzed zones of these vegetables is below the limits authorized by the European Union.[11]

Optimal spectral region and spectral repeatability

Before developing the models, the NIRS analysis of summer squashes had to be optimized in order to obtain a representative and quality spectrum per fruit or per zone, which in turn enable to obtain robust models when defining their safety and quality characteristics and to assess their possible industrial use in the baby food industry.

The existence of noise in the spectrum was then evaluated (spectral range 1600-2400 nm for the two MEMS spectrometers and 910-1676 nm for the LFV instrument).[17] After this process, the spectral range between 2240–2400 nm was eliminated in the Phazir 2400 and in the MicroPhazir due to the high level of noise detected (Fig. 1). Thus, all the models subsequently developed with these instruments were designed using the spectral range 1600–2240 nm. In the case of the MicroNIR-1700 instrument, as shown in Fig. 1, it does not produce high noise levels when working between 910-1676 nm, and for this reason, the full spectral range of the instrument was used.

Table 4 shows the STD_{limit} values for the two strategies tested, using the three NIRS spectrometers ~~tested~~. It is clearly shown that the values given by the STD_{limit} were lower when only 4 spectra per fruit were taken in one particular zone; in this case, the peduncle zone was chosen because of the greater accumulation of nitrates. As a result, to determine the destination of the analyzed fruit, for the two types of summer squash and NIRS instruments ~~tested~~, it would be enough to perform the NIRS and laboratory analysis only in this zone.

As can be seen, the lowest STD_{limit} values were obtained with the MicroNIR-1700 instrument for the two strategies and types of summer squashes ~~tested~~, while the results were reasonably similar in the two MEMS devices compared.

Once the RMS value did not exceed the value of the STD_{limit}, the spectra were then averaged.

The calculation of the RMS statistic is of extreme importance because it aims to ensure the spectral repeatability, which is essential for obtaining high quality spectral data, and therefore constitutes an essential step in obtaining robust equations.

No values for this statistic have been found in the scientific literature for summer squashes analyzed either whole or in zones on the vine, although the RMS statistic is extremely useful for obtaining representative spectral libraries of this vegetable, when they are analyzed on the plant.

Spectral properties

The ~~typical~~ mean log (1/R) spectra, together with the most relevant absorption bands for intact summer squashes scanned with Phazir 2400, MicroPhazir and MicroNIR-1700 are shown in Fig. 2.

In the 1600-2240 nm wavelength region for the two MEMS spectrometers tested, the major absorption peak at around 1920 nm is mainly related to water absorption, while there is another peak at 1780 nm, related to the first overtone of C-H stretching bonds.[28]

The mean spectrum obtained with MicroNIR-1700 shows a peak at 1450 nm related to the first overtone of the O-H group, as well as to the N-H stretch first overtone.[28] Moreover, the peak corresponding to the second overtone of O-H group can be seen at 970 nm.[29] Another peak can also be observed at approximately 1170 nm, which is linked to the second overtone of the C-H groups.[28]

Choice of the best handheld, portable NIRS instrument for *in situ* morphological, safety and quality determinations in summer squashes

Table 5 shows the statistics for the best calibration models obtained to predict the parameters studied using the three instruments tested. In order to compare the three spectrometers ~~tested~~, the calibration models for the different parameters in the study were carried out without eliminating the physical-chemical outliers during their development, which means that the values for mean, range and SD for each parameter are the same (Table 2).

Once the calibration models for the analyzed parameters were developed for each of the instruments tested, the SECV statistic values obtained for each parameter ~~in the study~~ were compared. As can be seen in Table 6, the SECV values corresponding to the weight and nitrate content parameters obtained with the MicroNIR-1700 are significantly lower ($P < 0.05$) than for the other two instruments used. As regards length, the SECV values obtained with the Phazir 2400 and MicroNIR-1700 are significantly lower ($P < 0.05$) than those obtained with the MicroPhazir. For the rest of the parameters ~~analyzed~~, no significant differences ($P > 0.05$) were found between the SECV values for the predictive models with the three instruments ~~tested~~.

In view of the results obtained, the MicroNIR-1700 instrument therefore, appears to be the most suitable for the analysis of the morphological, safety and quality parameters in summer squashes *in situ,* directly on the plant.

New calibration models for predicting morphological, safety and quality parameters in summer squash and external validation

After comparison of the spectrometers ~~tested~~ using the same number of samples, new calibration models with the sets C3 and C4 were constructed, but this time eliminating physical-chemical outlier samples if necessary; only the MicroNIR-1700 spectrometer was used for this purpose. The calibration statistics for the best models are shown in Table 7.

In the case of the morphological parameters, the models developed for the parameters of weight and equatorial diameter showed a predictive capacity ($r^2_{cv} = 0.84$, $RPD_{cv} = 2.49$) which could be considered good for both parameters, following the interpretation of the coefficient of determination values proposed by Shenk and Westerhaus[19] and Williams[24] while Nicolaï *et al.*[1] state that a RPD_{cv} value of between 2 and 2.5 indicates that coarse quantitative predictions are possible. As for length, the predictive capacity of the developed model (r^2_{cv} = 0.72; RPD_{cv} = 1.87) can be considered good,[19,24] while in Nicolaï *et al.*[1], the RPD_{cv} between 1.5 and 2 means that the model can discriminate between low and high values of the response variable.

The results obtained in this research are similar to those obtained by Sánchez *et al.*[30] for the prediction of the morphological parameters of weight (RPD_{cv} = 2.88), length (RPD_{cv} = 2.42) and equatorial diameter (RPD_{cv} = 2.26), using the Phazir 2400 in reflectance mode in the spectral range of 1600-2400 nm.

The satisfactory predictive capacity obtained for these morphological parameters using the handheld instrument MicroNIR-1700 is associated to the correlation between the size of the product and its water content. The absorption level of the light is highly dependent on the variation of these parameters, so it is possible to correlate the NIR signal with morphological parameters.[31]

As regards the determination of nitrate content, the model's predictive capacity (r^2_{cv} = 0.68; RPD_{cv} = 1.78) allows to discriminate between high, medium and low values, following the guidelines of Shenk and Westerhaus[19] and Williams,[24] and between high and low values according to the RPD_{cv} values suggested by Nicolaï *et al.*[1]

As far as measuring this parameter with NIRS technology, Sánchez *et al.*[9] obtained predictive capacity models (RPD_{cv} = 1.91) similar to the one obtained here (RPD_{cv} = 1.78) with the MicroNIR-1700 instrument.

Predicting the content in dry matter and soluble solids in summer squashes is extremely important in order to decide on the optimum moment for harvesting. The predictive capacity of the models for these parameters allows to differentiate between high, medium and low values for dry matter and between high and low values for SSC.[19,24]

As Fearn[32] points out, while the r^2_{cv} statistic can be a useful measure of the performance of a calibration, it does have its limitations. One major constraint is its dependence on the range of values—and on the SD of the reference values—of the calibration set. This would account for the lower r^2_{cv} values recorded here for both parameters, due to the reduced SD values shown.

Dardenne[33] and Fearn[32] have also shown that the RPD_{cv} statistic is equal to $1/\sqrt{1-r_{cv}^2}$ and depends to the same degree as r^2_{cv} on the range and SD of the data in the calibration set. This view is borne out by the results obtained here (Table 7), which indicate a close match between the highest and lowest r^2_{cv} values and RPD_{cv} values for the parameters tested.

Sánchez *et al.*[9] measuring dry matter and SSC in summer squashes obtained RPD_{cv} values of 1.75 and 1.56, respectively, using the Phazir 2400. The predictive capacity of both models is slightly higher than that obtained here, because the authors were able to use calibration sets with more variability.

Validations of the best calibration models obtained with calibration sets C3 and C4 were performed using the sets V3 and V4, respectively (Fig. 3). It is important to point out that some samples (N = 4 samples for weight; N = 2 samples for length; N = 3 equatorial diameter; N = 6 samples for nitrate content, N = 4 samples for dry matter, N = 1 sample for SSC), which were initially part of the V3 and V4 validation sets, were eliminated before the validation procedure was carried out with the calibration models developed with the MicroNIR-1700 instrument, due to they were hardly represented in the calibration sets with which the predictive models were finally designed.

Likewise, as regards the prediction of nitrate content, 4 samples which had a lower nitrate content (values below 103 mg kg^{-1}) were predicted by the models assigning them negative values for this parameter. However, the predictive NIRS values for these samples were shown as zero in Fig. 3.

After studying the results of the validation models, it can be affirmed that the standard error of prediction (SEP) values obtained are comparable to those from SECV, for the parameters tested. It is confirmed that the SECV is a good estimator of the SEP.[34]

According to the validation protocol established by Windham *et al.*[27] and once the results shown in Fig. 3 were analyzed, the models constructed for predicting all the morphological parameters analyzed, and also for the prediction of dry matter in intact summer squashes, met the validation requirements in terms of the coefficient of determination for prediction, r^2_p ($r^2_p > 0.6$) and both the standard error of prediction corrected for bias ($SEP_{(c)}$) and the bias were within confidence limits: the models thus ensure accurate prediction, and can be applied routinely. For the parameters nitrate content and SSC, it should be stressed that the $SEP_{(c)}$ and bias lay within the confidence limits, and although, r^2_p values did not attain the recommended minimum value (r^2_p =0.55 and 0.57 for nitrate content and SSC, respectively), they were close. Therefore, the results obtained suggest that the NIRS models developed can be regarded as a useful preliminary trial for obtaining accurate on-vine morphological, safety and quality predictions for intact summer squashes.

Furthermore, the external validation results obtained in this research for the morphological parameters of weight (RPD_p = 3.09), length (RPD_p = 2.37) and equatorial diameter (RPD_p = 3.10) are superior to those reported by Sánchez *et al.*[30] (RPD_p = 2.49, RPD_p = 1.59 and RPD_p = 1.67 for the three parameters mentioned above, respectively). For the nitrate content, the external validation value of RPD statistic (RPD_p = 1.60) is slightly lower than that obtained by Sánchez *et al.*[9] (RPD_p = 1.93) while for dry matter and SSC, the RPD_p values here obtained (1.52 and 1.84, respectively) were higher than those reported by the authors cited (RPD_p = 1.32 for dry matter; RPD_p = 1.22 for SSC).

CONCLUSIONS

The results obtained suggest that the greatest accumulation of nitrates in summer squashes takes place in the peduncle zone, and it is this area which must be analyzed to determine the destination of the harvested product. In addition, the summer squashes harvested at the end of the harvesting time should be the ones which are destined for baby food production, since they have nitrate values of below 200 mg kg^{-1}.

The findings also confirm the expectations raised that NIRS technology can enable intact summer squashes to be selectively harvested according to their morphological, safety and quality characteristics and to establish their industrial destination non-destructively. Additionally, the three NIRS instruments tested provided a similar level of accuracy for the measurement of equatorial diameter, dry matter and SSC. However, for weight, length and nitrate content, significantly more accurate models were obtained with the LVF instrument. The MicroNIR-1700 instrument is therefore, the most suitable for measuring, *in situ*, the morphology, safety and quality of summer squash.

ACKNOWLEDGEMENTS

This research was carried out under the Research Project 'Quality determination of summer squashes grown on an open-air plantation in Santaella (Córdoba)', funded by Gelagri Ibérica, S.L. The authors wish to thank to Ms. María Carmen Fernández for her technical assistance. The authors have no conflict of interest to declare.

REFERENCES

1. Nicolaï BM, Beullens K, Bobelyn E, Peirs A, Saeys W, Theron KI and Lammertyn J, Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review. *Postharvest Biol Technol* **46**:99 – 118 (2007).

2. Saranwong S and Kawano S, Applications to agricultural and marine products: fruits and vegetables, in *Near-infrared Spectroscopy in Food Science and Technology*, ed. by Ozaki Y, McClure WF and Christy AA. Wiley, Hoboken, NJ, pp. 219 – 242 (2007).

3. Sánchez MT and Pérez-Marín D, Non-destructive measurement of fruit quality by NIR spectroscopy, in *Advances in Post-Harvest Treatments and Fruit Quality and Safety* ed. by Vázquez M and Ramírez JA. Nova Science Publishers, Inc., Hauppauge, NY, pp. 101 – 163 (2011).

4. Magwaza LS, Opara UL, Nieuwoudt H, Cronje PJR, Saeys W and Nicolaï BM, NIR spectroscopy applications for internal and external quality analysis of citrus fruit—a review. *Food Bioprocess Tech* **5**:425 – 444 (2012).

5. Teixeira dos Santos CA, Lopo M, Páscoa RNMJ and Lopes JA, A review on the applications of portable near-infrared spectrometers in the agro-food industry. *App Spectrosc* **67**:1215 – 1233 (2013).

6. Yan H and Siesler HW, Hand-held near-infrared spectrometers: State-of-the-art instrumentation and practical applications. *NIR news* **29**:8 – 12 (2018).

7. De la Roza-Delgado B, Garrido-Varo A, Soldado A, Arrojo AG, Valdés MC, Maroto F and Pérez-Marín D, Matching portable NIRS instruments for *in situ* monitoring indicators of milk composition. *Food Control* **76**:74 – 81 (2017).

8. Sánchez MT, De La Haba MJ, Benítez-López M, Fernández-Novales J, Garrido-Varo A and Pérez-Marín D, Non-destructive characterization and quality control of intact strawberries based on NIR spectral data. *J Food Eng* **110**:102 – 108 (2012).

9. Sánchez MT, Pérez-Marín D, Torres I, Gil B, Garrido-Varo A and De la Haba MJ, Use of NIRS technology for on-vine measurement of nitrate content and other internal quality parameters in intact summer squash for baby food production. *Postharvest Biol Technol* **125**:122 – 128 (2017).

10. Torres I, Pérez-Marín D, De la Haba MJ and Sánchez MT, Developing universal models for the prediction of the physical quality in citrus fruits analysed on-tree using portable NIRS sensors. *Biosyst Eng* **153**:140 – 148 (2016).

11. Official Journal of the European Union (OJEU), *Commission Regulation (EC) No 1258/2011 of 2 December 2011 Amending Regulation (EC) No 1881/2006 As Regards Maximum Levels for Nitrates in Foodstuffs*. OJ L 320/15-17, 3.12.2011 (2011).

12. Tamme T, Reinik M, Roasto M, Juhkam K, Tenno T and Kiis A, Nitrates and nitrites in vegetables and vegetable-based products and their intakes by the Estonian population. *Food Addit Contam* **23**: 355 – 361 (2006).

13. Santamaria P, Elia A, Serio F and Todaro E, A survey of nitrate and oxalate content in retail fresh vegetables. *J Sci Food Agric* **79**:1882 – 1888 (1999).

14. Peiris KHS, Dull GG, Leffler RG and Kays SJ, Spatial variability of soluble solids or dry-matter content within individual fruits, bulbs, or tubers: implications for the development and use of NIR spectrometric techniques. *HortScience* **34**:114 – 118 (1999).

15. Long RL and Walsh KB, Limitations to the measurement of intact melon total soluble solids using near infrared spectroscopy. *Aust J Agr Res* **57**:403 – 410 (2006).

16. ISI, *The Complete Software Solution Using a Single Screen for Routine Analysis, Robust Calibration, and Networking. Manual, FOSS NIRSystems/TECATOR*. Infrasoft International, LLC, Silver Spring, MD (2000).

17. Hruschka WR, Data analysis: Wavelength selection methods, in *Near-Infrared Technology in the Agricultural and Food Industries*, ed. by Williams PC and Norris KH. American Association of Cereal Chemists, Inc., St. Paul, MN, pp. 39 – 58 (2001).

18. Shenk JS and Westerhaus MO, *Routine Operation, Calibration, Development and Network System Management Manual*. NIRSystems, Silver Spring, MD (1995).

19. Shenk JS and Westerhaus MO, Calibration: the ISI way, in *Near Infrared Spectroscopy: The Future Waves*, ed. by Davies AMC and Williams PC. NIR Publications, Chichester, pp. 198 – 202 (1996).

20. Martínez ML, Garrido A, De Pedro EJ and Sánchez L, Effect of sample heterogeneity on NIR meat analysis: the use of the RMS statistic. *J Near Infrared Spectrosc* **6**:313 – 320 (1998).

21. Shenk JS and Westerhaus MO, Population structuring of near infrared spectra and modified partial least squares regression. *Crop Sci* **31**:1548 – 1555 (1991).

22. Barnes RJ, Dhanoa MS and Lister SJ, Standard Normal Variate transformation and De-trending of near infrared diffuse reflectance spectra. *Appl Spectrosc* **43**:772 – 777 (1989).

23. Shenk JS and Westerhaus MO, *Analysis of Agriculture and Food Products by Near Infrared Reflectance Spectroscopy*. NIRSystems, Silver Spring, MD (1995).

24. Williams PC, Implementation of near-infrared technology, in *Near Infrared Technology in the Agricultural and Food Industries*, ed. by Williams PC and

Norris KH. American Association of Cereal Chemists, Inc., St Paul, MN, pp. 145 – 171 (2001).

25. Massart DL, Vandeginste BGM, Deming SM, Michott Y and Kaufman L, *Chemometrics: A Textbook. (Data Handling in Science and Technology 2)*. Elsevier Science, Amsterdam (1988).

26. Naes T, Isaksson T, Fearn T and Davies A, *A User-Friendly Guide to Multivariate Calibration and Classification*. NIR Publications, Chichester (2002).

27. Windham WR, Mertens DR and Barton II FE, Protocol for NIRS calibration: sample selection and equation development and validation, in *Near Infrared Spectroscopy (NIRS): Analysis of Forage Quality (USDA-ARS Agriculture Handbook No. 643)*, ed. by Martens GC, Shenk JS and Barton II FE. US Government Printing Office, Washington, DC, pp. 96 – 103 (1989).

28. Shenk JS, Workman JJJr and Westerhaus MO, Application of NIR spectroscopy to agricultural products, in *Handbook of Near-Infrared Analysis*, ed. by Burns DA and Ciurczak E. Marcel Dekker, Basel, pp. 347 – 386 (2008).

29. Osborne BG, Fearn T and Hindle P, *Practical NIR Spectroscopy with Applications in Food and Beverage Analysis*. Longman Scientific and Technical, London (1993).

30. Sánchez MT, Torres I, Gil B, Pérez-Marín D, Garrido-Varo A and De la Haba MJ, In situ determination of external quality parameters in intact summer squash using near-infrared reflectance spectroscopy. *Acta Hort* **1194**:1259 – 1264 (2018).

31. Pérez-Marín D, Sánchez MT, Paz P, Soriano MA, Guerrero JE and Garrido-Varo A, Non-destructive determination of quality parameters in nectarines during on-tree ripening and postharvest storage. *Postharvest Biol Technol* **152**:180 – 188 (2009).

32. Fearn T, The overuse of R^2. *NIR news* **25**:32 – 32 (2014).

33. Dardenne P, Some considerations about NIR spectroscopy: Closing speech at NIR-2009. *NIR news* **21**:8 – 9 (2010).

34. Shenk JS, Westerhaus MO and Abrams SM, Protocol for NIR calibrations: monitoring analysis results and recalibration, in *Near Infrared Spectroscopy (NIRS): Analysis of Forage Quality (USDA-ARS Agriculture Handbook No. 643)*, ed. by Martens GC, Shenk JS and Barton II FE. US Government Printing Office, Washington, DC, pp. 104 – 110 (1989).

Table 1. Technical features of the spectrometers Phazir 2400, MicroPhazir and MicroNIR-1700.

Property	Instrument		
	Phazir 2400	MicroPhazir	MicroNIR-1700
Detector type	Single-element InGaAs detector	Single-element InGaAs detector	128-pixel InGaAs photodiode array
Dispersion element	MEMS	MEMS	LVF
Wavelength range (nm)	1600-2400	1600-2400	910-1676
Resolution (nm)	≈ 8	≈ 8	6.2
Sampling integration time (ms)	600	600	11
Weight (kg)	1.7	1.2	$64\cdot10^{-3}$
Analysis mode	Reflectance	Reflectance	Reflectance

Table 2. Number of samples (N), range, mean, SD, and coefficient of variation (CV) for the different calibration (C1, C2, C3 and C4) and validation (V3 and V4) sets.

	Parameters																	
	Weight (g)			Length (cm)			Equatorial diameter (mm)			Nitrates (mg kg^{-1})			Dry matter (% fw)			SSC (ºBrix)		
	C1	C3	V3	C1	C3	V3	C1	C3	V3	C2	C4	V4	C2	C4	V4	C2	C4	V4
N	217	163	54	217	163	54	217	163	54	407	306	101	407	306	101	407	306	101
Range	78.43-1746.49	78.43-1746.49	125.12-1135.89	12.83-43.50	12.83-43.50	16.60-43.00	28.02-89.58	28.02-89.58	30.48-83.95	18.50-1979.96	18.50-1979.96	55.92-1209.18	3.16-7.56	3.16-7.56	3.61-7.25	2.80-6.50	2.80-6.50	2.80-5.70
Mean	463.92	457.39	483.51	24.12	24.05	24.33	52.57	52.17	53.8	362.61	356.55	380.96	4.69	4.68	4.71	4.13	4.13	4.14
SD	302.38	300.09	311.21	5.63	5.67	5.57	14.29	14.16	14.75	292.67	299.06	272.99	0.72	0.68	0.83	0.47	0.46	0.50
CV	65.18	65.61	64.36	23.34	23.58	22.89	27.18	27.14	27.42	80.71	83.88	71.66	15.38	14.53	17.62	11.38	11.14	12.08

Table 3. Evolution of nitrate content in the three zones analyzed during the harvest period in summer squashes weighing over 400 g.

Harvest date	Nitrate content (mg kg^{-1})		
	Peduncle zone	Equatorial zone	Stylar Zone
05/17/2017	1056.55 (502.58)(a)	758.96 (359.49)(b)	827.43 (330.33)(a,b)
05/22/2017	520.87 (96.49)(e)	405.10 (34.64)(f)	464.98 (64.30)(e,f)
05/31/2017	750.77 (249.37)(c)	624.78 (180.81)(d)	694.48 (208.89)(c,d)
06/05/2017	516.51 (225.14)(e)	452.77 (164.12)(f)	500.39 (193.19)(e,f)
06/12/2017	343.34 (157.49)(g)	283.95 (106.10)(h)	320.45 (109.02)(g,h)
06/20/2017	193.85 (109.82)(i,k)	175.56 (92.51)(j,l)	201.07 (114.18)(i,j,k,l)
06/26/2017	240.00 (133.85)(i)	213.71 (125.67)(j)	225.39 (127.70)(i,j)
07/06/2017	181.73 (43.09)(i,k)	153.39 (49.86)(j,l)	172.74 (61.87)(i,j,k,l)
07/13/2017	91.25 (68.94)(k)	66.00 (50.91)(l)	84.25 (66.11)(k,l)

Standard deviation in brackets.

The same letter indicates homogeneous group established by ANOVA ($P < 0.05$).

Table 4. STD_{limit} (μlog(1/R) of the RMS statistic for summer squashes analyzed on-vine.

Spectrometer	Summer squash			
	Weight > 400 g		Weight < 400 g	
	Strategy I 4 spectra	Strategy II 12 spectra	Strategy I 4 spectra	Strategy II 12 spectra
Phazir 2400	53,822	65,290	52,659	62,893
MicroPhazir	44,304	61,560	49,177	63,818
MicroNIR-1700	29,205	29,711	47,533	51,784

Table 5. Calibration statistics for NIR-based models for predicting morphological, safety and quality parameters in intact summer squash.

Parameter	Instrument	Math treatment	SECV	r^2_{cv}	RPD_{cv}
Weight (g)	Phazir 2400	1,5,5,1	155.91	0.73	1.94
	MicroPhazir	2,5,5,1	161.07	0.72	1.88
	MicroNIR-1700	2,5,5,1	142.48	0.78	2.12
Length (cm)	Phazir 2400	1,5,5,1	3.31	0.65	1.70
	MicroPhazir	1,5,5,1	3.40	0.64	1.66
	MicroNIR-1700	1,5,5,1	3.11	0.69	1.81
Equatorial diameter (mm)	Phazir 2400	1,5,5,1	6.34	0.80	2.25
	MicroPhazir	1,5,5,1	6.70	0.78	2.13
	MicroNIR-1700	2,5,5,1	6.22	0.81	2.30
Nitrate content (mg kg^{-1})	Phazir 2400	2,5,5,1	240.03	0.33	1.22
	MicroPhazir	1,5,5,1	226.02	0.40	1.29
	MicroNIR-1700	1,5,5,1	198.07	0.54	1.48
Dry matter (% fw)	Phazir 2400	2,5,5,1	0.53	0.46	1.36
	MicroPhazir	2,5,5,1	0.53	0.47	1.36
	MicroNIR-1700	1,5,5,1	0.51	0.50	1.41
SSC (ºBrix)	Phazir 2400	1,5,5,1	0.36	0.43	1.31
	MicroPhazir	1,5,5,1	0.35	0.45	1.34
	MicroNIR-1700	1,5,5,1	0.33	0.50	1.42

Table 6. Comparison between SECV values obtained for the best models for predicting the morphological, safety and quality parameters of summer squashes using the MEMS and LVF spectrometers tested; Fisher test ($P < 0.05$).

Parameter	SECV Phazir 2400	SECV MicroPhazir	SECV MicroNIR-1700	$SECV_{min}$	$SECV_{min} \cdot \sqrt{F_{critical}}$
Weight (g)*	155.91	161.07	142.48	142.48	154.63
Length (cm)*	3.31	3.40	3.11	3.11	3.37
Equatorial diameter (mm)	6.34	6.70	6.22	6.22	6.96
Nitrate content (mg kg^{-1})*	240.03	226.02	198.07	198.07	214.94
Dry matter (% fw)	0.53	0.53	0.51	0.51	0.55
SSC (ºBrix)	0.36	0.35	0.33	0.33	0.36

*: Significant differences ($P < 0.05$) between the SECV values obtained.

Table 7. Calibration statistics for best NIR-based models for predicting morphological, safety and quality parameters in intact summer squashes using the MicroNIR-1700 instrument.

Parameter	Math treatment	N	Range	Mean	SD	SECV	r^2_{cv}	RPD_{cv}
Weight (g)	1,5,5,1	154	78.43-1388.16	420.70	250.33	100.65	0.84	2.49
Length (cm)	1,5,5,1	155	12.83-40.00	23.62	5.22	2.79	0.72	1.87
Equatorial diameter (mm)	2,5,5,1	156	28.02-89.58	52.53	13.87	5.57	0.84	2.49
Nitrate content (mg kg^{-1})	2,5,5,1	294	18.50-1219.73	325.72	251.12	141.32	0.68	1.78
Dry matter (% fw)	2,5,5,1	297	3.16-7.51	4.64	0.62	0.42	0.54	1.48
SSC (ºBrix)	2,5,5,1	300	2.80-5.20	4.12	0.43	0.31	0.47	1.39

Figure 1. D_1 log (1/R) spectra for summer squash. Instruments: a) Phazir 2400, b) MicroPhazir and c) MicroNIR-1700.

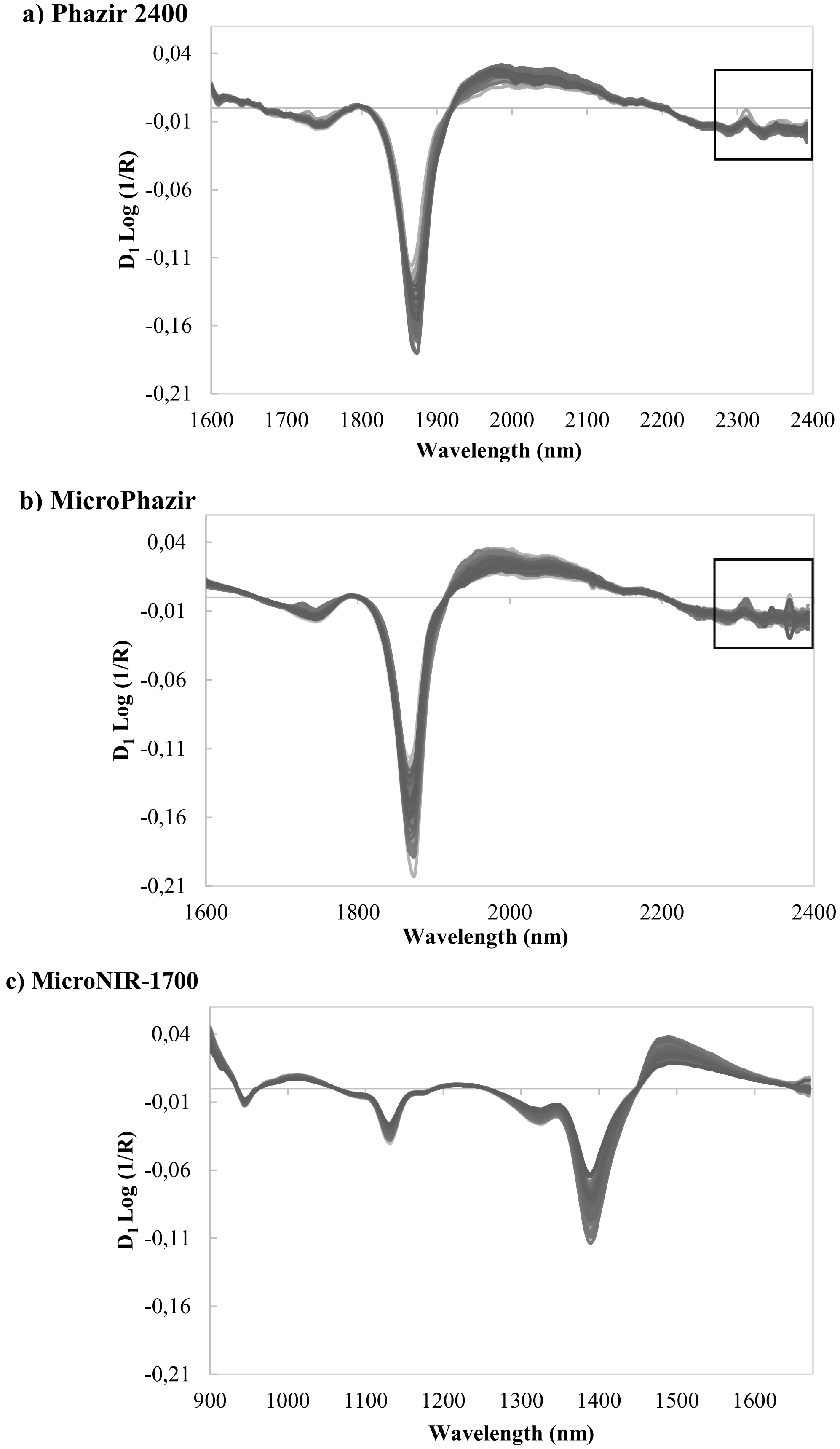

Figure 2. Mean spectra for summer squash. Instruments: Phazir 2400, MicroPhazir and MicroNIR-1700.

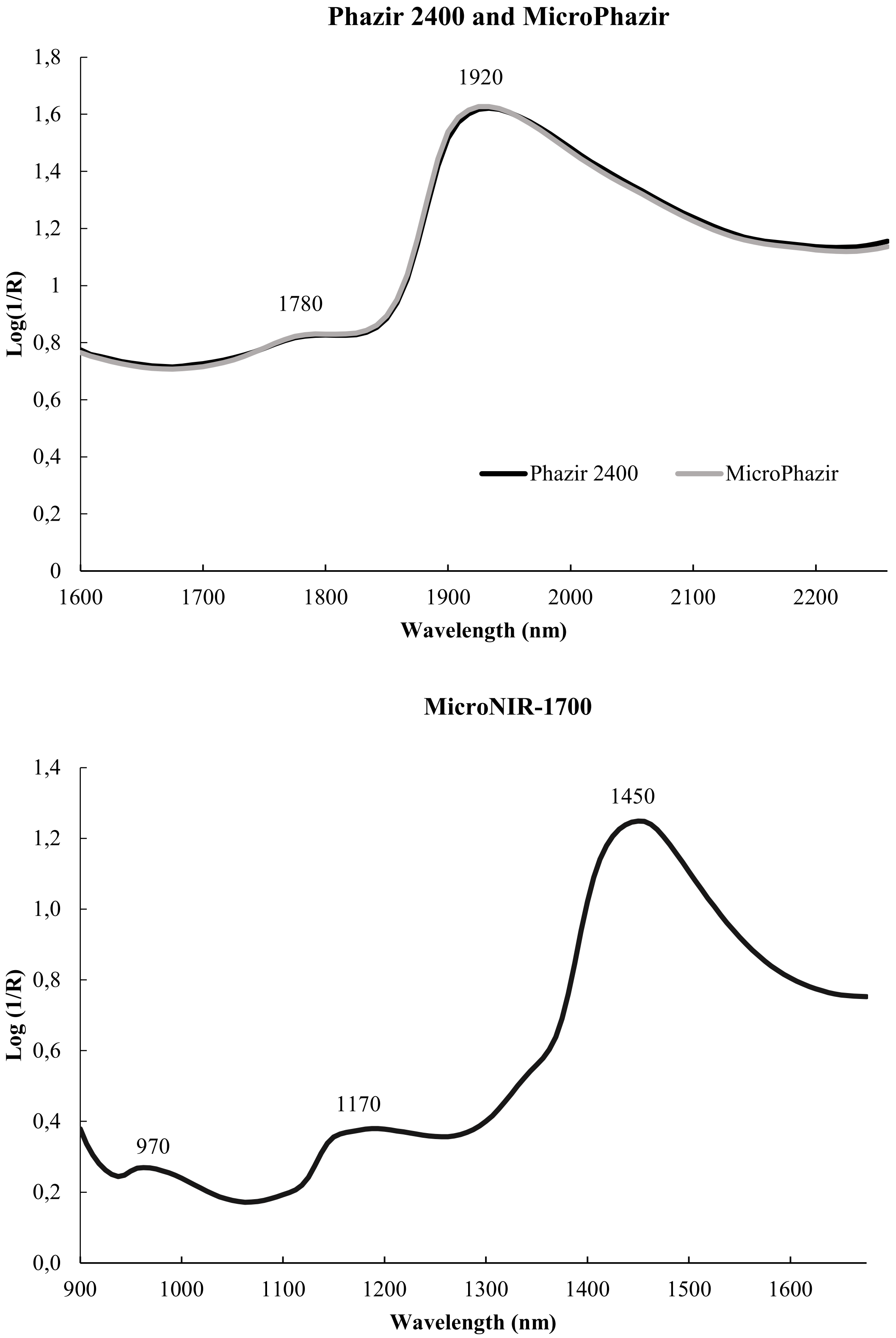

Figure 3. Reference *versus* NIR-predicted data for the validation sets. N, number of samples for the validation set; SEP, standard error of prediction; SEP(c), standard error of prediction corrected for bias; r^2_p, coefficient of determination for prediction; RPD_p, residual predictive deviation for prediction; fw, fresh weight