PRE-HARVEST SCREENING ON-VINE OF SPINACH QUALITY AND SAFETY USING NIRS TECHNOLOGY

Dolores Pérez-Marín [a,*], Irina Torres [b], José-Antonio Entrenas [b], Miguel Vega [b], María-Teresa Sánchez [b,*]

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

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

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

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

Abstract

The study sought to perform a non-destructive and *in-situ* quality evaluation of spinach plants using near infrared (NIR) spectroscopy in order to establish its suitability for different uses once harvested. Modified partial least square (MPLS) regression models using NIR spectra of intact spinach leaves were developed for nitrate, ascorbic acid and soluble solid contents. The residual predictive deviation (RPD) values were 1.29, 1.21 and 2.54 for nitrate, ascorbic acid and soluble solid contents, respectively. Later, this predictive capacity increased for nitrate content (RPDcv = 1.63) when new models were developed, taking into account the influence on the robustness of the model exercised by the simultaneity between the NIR and laboratory analyses. Subsequently, using partial least squares discriminant analysis (PLS-DA), the ability of NIRS technology to classify spinach as a function of nitrate content was tested. PLS-DA yielded percentages of correctly classified samples ranging from 73.08-76.92% for the class 'spinach able to be used fresh' to 85.71-73.08% for the class 'preserved, deep-frozen or frozen spinach, both for unbalanced and balanced models respectively, based on N-H signal associated with proteins. Overall, the data supports the capability of NIR spectroscopy to establish the final destination of the production of spinach analysed on the plant, as a screening tool for important safety and quality parameters.

Keywords: Spinach; Safety; Quality; Portable NIRS; In-situ analysis

1. Introduction

Spinach (*Spinacea oleracea* L.) is a green, leafy vegetable, with a shiny, uniform appearance. When sold, either to industry (for preserving, freezing or deep-freezing) or as a fresh product on the market, it is essential for the leaves to look freshly-picked and tender.

Spinach has an extremely high water content (around 92%) and is very low in carbohydrates and fats. It is also a rich source of vegetable protein, with, like other vegetables, high fibre content [1].

Spinach is also noted for its relatively high content of bioactive substances, including vitamin C (ascorbic acid), which is a powerful antioxidant that participates in the scavenging of the reactive oxygen species, regenerating tocopherols from their radical forms [2, 3]. However, one ~~additional~~ drawback is that they accumulate substances which are harmful for the human organism, such as nitrates [4].

Over recent years, consumers have become increasingly aware of the presence of nitrates in foods, among which are vegetables, since nitrates are a serious threat to human health, due to the conversion of nitrate to nitrite, which may produce methemoglobin due to the oxidation of Fe+2 in haemoglobin [5]. The impaired capacity of methemoglobin to deliver oxygen to tissues may lead to severe toxic effects and may even prove fatal where methemoglobin accounts for over 70% of total haemoglobin, something which affects infants and very young children almost exclusively [6]. Similarly, a number of studies have highlighted a possible link between nitrate exposure and childhood type 1 insulin-dependent diabetes mellitus [7]. Furthermore, nitrite may react with secondary amines (HNR_2), which occur in many foods, to form nitrosamines. These substances are highly carcinogenic [8].

However, consumers are also well aware that eating vegetables with a high content of antioxidants, such as ascorbic acid, is beneficial for their health. The World Health Organization [9] showed that iron deficiency anemia is one of the most common nutritional disorders which has a major effect on both health and the economy. Main cause of anemia is not only low iron intake but also poor iron absorption [10]. This global health problem can be addressed by improving the dietary iron bioavailability, which can be altered by various components present in the food which can either enhance or inhibit iron absorption. Thus, when iron is present along with ascorbic acid, the absorption of iron has been shown to increase even in the presence of inhibitors [11].

In the case of spinach and in response to this growing public concern about nitrates, the European Union passed Commission Regulation (EC) No 1258/2011 of 2 December 2011 setting maximum levels for nitrates in this leafy vegetable [12]. Thus, the maximum level for nitrates was set for preserved, deep-frozen or frozen spinach at 2,000 mg NO_3/kg and for fresh spinach at 3,500 mg NO_3/kg.

All this has prompted greater attention to spinach safety and quality concerns. Nitrate accumulation, ascorbic and soluble solid contents in spinach depend not only on genotypic characteristics, but also on a number of other factors, including cultural practices, harvesting date and postharvest handling practices [13, 14]. As a result, producers are increasingly anxious to provide consumers with assurances regarding the safety, quality and provenance of this product.

The nitrate content in spinach when harvested is in fact key to determining the final destination of the harvested product. The ascorbic acid content is equally important because of the close relationship between this acid and the bioavailability of iron, and also the soluble solid content, which in turn is linked to the vegetable's quality and shelf-life.

92 NIR spectroscopy in conjunction with the application of multivariate analysis
93 strategies is an appropriate non-destructive technology for the study of chemical
94 constituents of vegetables at field level. This technology represents a marked change from
95 the conventional analytical methods, because a single spectrum allows the simultaneous
96 characterization of different chemical properties, in a matter of seconds and without
97 sample preparation, thus allowing real-time decision making. In the field, it has become
98 easier to use this technology by the development in recent years of compact, portable
99 hand-held instruments which make it possible to measure the quality and safety
100 parameters of vegetables directly on the plant, thus allowing the product to be instantly
101 analysed.

102 Several authors have shown the feasibility of using NIRS technology for the non-
103 destructive measurement of nitrate content in various fruits and vegetables, including
104 Japanese radishes [15], the leaf stalk of Qing gin cai [16], pineapple [17] and summer
105 squash [18]. In spinach, Xue and Yang [19] and Itoh et al. [20], are the only ones to have
106 carried out this analysis, although these authors used NIRS instruments with optical
107 performance and wavelength ranges different to those used here.

108 As regards ascorbic acid content, no references have been found where this
109 parameter has been measured in spinach using NIRS. However, some authors [21-25]
110 have shown how NIRS technology can be used to measure ascorbic acid in apples,
111 zucchini, oranges, potatoes and peppers.

112 In the case of soluble solid content (SSC), no references have been found for
113 measuring this parameter with NIRS in spinach, although several review works indicate
114 that NIRS technology is a viable means of measuring this parameter in fruit and
115 vegetables [26-28].

Taking into account the possibilities of using handheld NIRS sensors to optimize harvesting times and enable the staggered harvesting of spinach for greater quality and safety, thus allowing certain harvested spinach to be used either in the production of baby foods, industrially processed as preserved, deep-frozen or frozen spinach or for fresh consumption, this study sought to assess the feasibility of using NIR spectroscopy~~,~~ to characterize the variations in internal safety and quality—particularly nitrate, ascorbic and soluble solid contents—in intact spinach during on-vine ripening using a low-cost, miniaturized, handheld, near-infrared device based on the MEMS system.

2. Material and methods

2.1. Sampling

A total of 128 samples of spinach plants (*Spinacia oleracea* L, cv. 'Solomon' (62 samples), 'Novico' (13 samples), 'Meerkat' (10 samples), and 'Gorilla' (43 samples), grown in an open-air plantation in the province of Córdoba (Spain), were harvested between January and March 2017.

2.2. Reference data

Nitrate content (mg NO_3/kg) was measured using an RQFlex reflectometer (Merck, Darmstadt, Germany) [29]. The reflectometer which measures the colour intensity of Reflectoquant ® test strips (Merck, Darmstadt, Germany) is based on a colorimetric method. For NO_3^- analysis, the spinach leaves were cut into very small pieces and liquified. Next, 5 ml of spinach juice was mixed in a blender with different quantities

of deionised water, depending on NO_3^- concentrations. After that, the solution was filtered using a coffee filter and left to settle for 5 min. Subsequently, a test strip was dipped in the supernatant for 2 s, and then the colour was allowed to develop for 1 min. The test strip was then inserted into the reflectometer and the amount of light reflected from the test strip was measured and converted to a concentration by a standard calibration previously introduced into the equipment using a bar-coded plastic strip. The dilution factor was also taken into consideration.

Ascorbic acid content (mg/100g) was also measured using a reflection photometer (RQflex 10, Merck, Merck, Darmstadt, Germany) [30]. For the ascorbic acid analysis, the analysis procedures were the same as those for nitrate content with the exception that samples containing more than 450 mg/L of ascorbic acid were diluted with oxalic acid solution 1%. The dilution factor was also taken into consideration.

SSC (ºBrix) was measured as the refractometer reading for the spinach juice, using a temperature-compensated digital Abbé-type refractometer (model B, Zeiss, Oberkochen, Würt, Germany).

2.3. Spectral data acquisition

The NIR spectra of the spinach leaves were collected in reflectance mode (log 1/R) using a handheld MEMS instrument (Phazir 2400, Polychromix, Inc., Wilmington, MA, USA). The Phazir 2400 is a compact, low-cost near-infrared analyser that incorporates all the essential components to deliver on-vine applications. The spectrophotometer 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 only around 4 mm^2. The sensor integration time was 600 ms. The MEMS device is equipped with quartz protection

166 to prevent the accumulation of dirt. Instrument performance was checked every 10 min,
167 following the diagnostic protocols provided by the manufacturer, and white reference
168 measurement was carried out using Spectralon as reference.

169 Four spectral measurements were made on each spinach leaf (distal and proximal,
170 on both sides (right and left) of the leaf blade relative to the main vein, on the adaxial
171 side). Because between 4-10 leaves were used for the chemical analyses of the parameters
172 of SSC, ascorbic acid and nitrate content when analysing each plant, to obtain a mean
173 spectrum for each spinach sample, a mean spectrum was first obtained from the four
174 spectra for each leaf, and then a mean spectrum was obtained from the four to ten mean
175 spectra for each sample.

176

177 *2.4. Spectral repeatability*

178

179 The spectral repeatability of intact spinach leaves was evaluated using the Root
180 Mean Squared (RMS) statistic. The RMS statistic is defined as the averaged root mean
181 square of differences between the different subsamples scanned at *n* wavelengths [31,
182 32]. This statistic indicates the similarity between different spectra of a single sample.

183 For each instrument and sample presentation form, the RMS for an individual
184 subsample (*j)* of the sample (*k*) can be calculated using the using the following expression:

185 $$RMS_{j,k} = \sqrt{\frac{\sum_{i=1}^{n} D_{ij}^2}{n}};\ D_{ij} = y_{ij} - \bar{y}_i$$

186 where y_{ij} is log (1/R) at wavelength *i* for subsample *j*, and $\bar{y}_i$ is log (1/R) at wavelength *i*
187 for the average spectrum of a sample *k*; *n* is the number of data points collected by the
188 instrument (here, 100 data points for the MEMS instrument). The RMS value obtained in
189 each case was multiplied by 10^6 to facilitate value management and processing.

To determine a cut-off value (RMS_{cutoff}) of each sample presentation form, the mean RMS was calculated along with standard deviation (STD) per sample according to the formulae provided by Martínez et al. [33].

$$STD_k = \sqrt{\sum_{j=1}^{N} (RMS_j)^2/(N-1)}$$

where N are the number of sub-samples.

A STD limit can be calculated for comparing the RMS values of subsamples, following the formula provided by Rosales [34], who demonstrated that the estimated value of the error variance, σ^2 for log (1/ R) $y_{i,j}$ is the corresponding to one-way ANOVA:

$$\sigma^2 = \frac{1}{n\,(N-1)} \sum_{i=1}^{n} \sum_{j=1}^{N} \left(y_{i,j} - \bar{y}_i\right)^2$$

This expression corresponds to the STD^2. The sum of squares for error (SSE) can be thus be expressed as:

$SSE = n\,(N\text{-}1)\,STD^2$

which approximately follows a χ^2 distribution,

$$\frac{n\,(N-1)\,(STD^2)}{\sigma_0^2} \sim \chi^2_{[n(N-1)]}$$

with σ_0^2 the parametric value of the error variance. For infinite degrees of freedom (> 100), χ^2 tends to a normal distribution. An STD limit can then be calculated for comparing the RMS values of subsamples, following the formula given by Rosales (1993).

$$STD_{limit} = 1.036 \sqrt{\sum_{k=1}^{k=m} STD_k^2/m} = 1.036\sqrt{\overline{STD^2}}$$

where STD is the standard deviation per sample and m is the number of samples. The value 1.036 corresponds to a probability level of 85%.

The STD_{limit} value was used to establish the RMS_{cutoff} for each product and analysis mode. Hence, the different sources of variation which might cause irregular spectra were controlled, since any spectra in a sample that were above this limit were eliminated, and recalculations were performed until all the values were below the RMS_{cutoff}. Then, the mean spectrum of each leaf was calculated.

To evaluate spectral repeatability, two alternatives were available: first, analysing ten leaves and taking two spectra in each of them, at the same point, and second, analysing twenty leaves and taking four spectra in each, at different points of the leaf.

2.5. Data analysis: definition of calibration and validation sets

Principal Component Analysis (PCA) was performed on a set of N = 128 samples in order to decompose and compress the data matrix. After PCA, the center of the spectral population was determined in order to detect outlier samples. The Mahalanobis distance (GH) was calculated between each sample and the center; samples with a GH value greater than 3 were considered outliers [31]. As spectral pre-treatments, the Standard Normal Variate (SNV) plus Detrending (DT) procedure [35] was used to remove the multiplicative interferences of scatter, and first derivative mathematical treatment (Norris derivative) was performed (1,5,5,1), where the first digit is the order of the derivative, the second is the gap over which the derivative is calculated, the third is the number of data points in a running average or smoothing and the fourth is the second smoothing [36, 37].

Once spectral outliers had been removed, a set consisting of 124 samples was used to develop the calibration models. The set was divided into two: a training set containing about 75% of the samples (N = 93 samples) and a test set containing the remaining 25% (N = 31 samples). These samples were selected following the method outlined by Shenk

235 and Westerhaus [38] using the CENTER algorithm included in the WinISI software
236 package to calculate the Global Mahalanobis distance (GH). The samples were ordered
237 on the basis of the Mahalanobis distance to the center of the population, and three of every
238 four were selected to be part of the calibration set.
239
240 *2.6. Chemometric tools*
241
242 The data were subjected to chemometric treatment using the WinISI software
243 package version 1.50 [37].
244 NIR calibration models for the prediction of quality parameters (nitrate content,
245 ascorbic content, and SSC) in intact spinach plants were constructed using modified
246 partial least squares (MPLS) regression [31], with subsequent cross-validation. The
247 calibration set was partitioned into 6 groups; each group was then validated using a
248 calibration derived from the other samples; finally, validation errors were combined to
249 obtain a standard error of cross-validation (SECV).
250 For each analytical parameter, different mathematical treatments were evaluated.
251 For scatter correction, the standard normal variate (SNV) and detrending (DT) methods
252 were tested [35]. Additionally, a total of four derivative mathematical treatments were
253 tested: 1,5,5,1; 2,5,5,1; 1,10,5,1 and 2,10,5,1.
254 The statistics used to select the best equations were: the coefficient of
255 determination for calibration (r^2_c), the standard error of calibration (SEC), the coefficient
256 of determination for cross calibration (r^2_{cv}), and the standard error of cross validation
257 (SECV). Furthermore, the Residual Predictive Deviation (RPD) for cross-validation was
258 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 [39].

The best-fitting equations obtained for the calibration set, as selected by statistical criteria, were subsequently subjected to external validation following the protocol outlined by Windham et al. [40].

After analysing the results obtained, and in order to test the influence of the simultaneity in time between the NIRS spectrum and the wet-chemistry analysis on the robustness of the model obtained for the prediction of nitrate content in intact spinach, new predictive models were designed for this parameter, dividing the initial total of 128 samples into 2 groups which represent two different analysis strategies.

- Strategy I. Group of samples 1 to 47. On the day the field samples arrived, the corresponding NIR spectra were taken. However, the reference analyses were carried out 24 hours later, and the samples were stored in refrigeration conditions during that time (4ºC, RH: 85%).

- Strategy II. Group of samples 48 to 128. The NIRS spectra were taken and the wet analysis performed 24 hours after the product arrived in the laboratory, and the samples were stored in refrigeration conditions until analysis (4ºC, RH: 85%).

The same signal pre-treatments and spectral region described earlier were used here for the development of the new quantitative models.

2.7. NIRS classification models

The design of models to classify spinach by its nitrate content, in order to evaluate the viability of using NIRS technology to determine the final destination of the harvested spinach (fresh consumption or preserved, deep-frozen or frozen product), for those

284 samples where the NIR and reference analyses were performed at the same time (Strategy
285 II), comprised two classification groups: 1) spinach which contained 0-2,000 mg/kg NO_3^-
286 (N = 45 samples) for preserved, deep-frozen or frozen spinach; 2) spinach with a nitrate
287 content between 2,000-3,500 mg/kg NO_3^- (N = 36 samples) for fresh spinach.

288 Next, the structure and spectral variability of the sample population was studied
289 in order to select the samples that would constitute the learning group. To do this, we
290 used the CENTER algorithm, which is included in the WinISI version 1.50 software. This
291 algorithm was applied separately to each of the two training groups. The mathematical
292 treatments SNV (Standard Normal Variate) and DT (Detrend) were applied to correct any
293 scattered radiation phenomena, together with the 1,5,5,1 derivation treatment [31].

294 After having ordered the sample set by spectral distances (from smallest to
295 greatest distance from the center), a structured selection of the external validation set (10
296 samples for each classification group), solely on the basis of spectral data [38].

297 Discriminant models were constructed to classify spinach by its NO_3^- content,
298 using PLS-DA for supervised classification. Specifically, the PLS2 algorithm was
299 applied, using the 'Discriminant Equations' option in the WINISI version 1.50 software
300 package [37].

301 All the models were constructed using six cross-validation groups. The same
302 signal pre-treatments and spectral regions described earlier for quantitative analysis were
303 used for qualitative model development.

304 The precision of the models obtained was evaluated using the percentage of
305 correctly-classified samples, both for the global model and for each class.

306 The difficulty involved in obtaining balanced learning groups in terms of the
307 number of samples per class or classification category led us to assess the influence of
308 this factor on the development of discriminant models. In this way, the results obtained

were contrasted with balanced and unbalanced classification models in terms of the number of samples per class.

The samples for the balanced groups were selected using the algorithm SELECT included in the WinISI II software package version 1.50 (Infrasoft International, Port Matilda, PA), which calculates spectral distances (Mahalanobis H), in order to detect samples whose spectrum is very similar to that of others in the population [38]. This algorithm enables spectral selection of a number of samples representative of the population as a whole, by calculating the 'NH' distance (Mahalanobis neighbour distance) between two spectra. An 'NH' of less than 0.6 implies that two spectra are too similar to each other ('neighbour'). After application of this algorithm, 26 samples of the category 'nitrate content between 0-2,000 mg/kg' were selected, thus making the number of samples of the two classes equal and allowing the classification models to be redesigned.

Next, the best classification models for each of the established types (unbalanced and balanced models) were selected and externally validated. In this case, an external validation procedure was carried out to measure the predictive capacity of the model using a sample group different to that used in the training of the model. In both models (unbalanced and balanced) 20 samples (10 per category) were selected in a structured way [38].

3. Results and discussion

3.1. Spectral repeatability

The collection of high quality spectra is crucial for the characterization of spinach plants by quality and safety characteristics and to assess its possible industrial use, as well as to construct discriminant classification models for the product depending on its possible use in the processing industry of fresh or processed vegetables (in this case, preserved, deep-frozen or frozen spinach). The RMS cut-off was calculated for the instrument MEMS used as shown in section 2.4.

For the first alternative tested (analysing ten leaves and taking two spectra in each of them, at the same point) the mean STD was 48,292 µlog (1/R), representing an RMS cut-off of 64,661 µlog (1/R). For the second alternative (analysing twenty leaves and taking four spectra in each, at different points of the leaf) the mean STD and the RMS cut-off were 118,693 µlog (1/R) and 128,437 µlog (1/R), respectively. As can be seen, the result obtained for $STD_{límit}$ in the test in which 10 samples (leaves) were used was lower than that obtained in the test which used 20 samples. This was to be expected, since the former corresponds to the analysis of 10 samples whose two subsamples were taken at the same point, while, in the latter case, 20 samples were analysed by taking the four subsamples at different points in the leaf. The repeatability results obtained therefore indicate that, in the second mode, the analysis reflects the heterogeneity of the leaf, although the reduced window of analysis presented by the spectrophotometer used must also be taken into account.

It was therefore decided to choose the second mode of analysis as the most suitable, that is, to collect 4 spectra per leaf. In this way, a more representative measurement of the analysed product was obtained. Once the RMS value did not exceed the value of the $STD_{límit}$, these spectra were then averaged in order to carry out both the quantitative and qualitative predictive models using the average spectrum of each sample.

The calculation of the RMS statistic is extremely important because it aims to ensure high spectral repeatability, which is essential to obtain 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 spinach analysed on the vine, although the RMS statistic is extremely useful to obtain representative spectral libraries of this vegetable, when analysed on the plant. In fact, this is the first research work to deal with measuring spectral repeatability in leafy vegetables.

3.2. Spectral features

Mean and standard deviation log (1/R) spectra for intact spinach leaves, captured by the instrument Phazir 2400, together with the most relevant absorption bands, are shown in Fig. 1.

In the NIR region between 1600-2400 nm, the highest absorption peak is at the 1920 nm wavelength, which corresponds ~~together with the 2200 nm valley,~~ to water. This was to be expected, since spinach is made up of 90% water [41]. Osborne et al. [42] showed that the peak in the wavelength 1780 nm was directly related to the first overtone of sugars. ~~The valley at 1670 nm could be related to the N-H groups [41]~~.

3.3. Descriptive data for NIR calibration and validation

Values for number of samples, range, mean, standard deviation (SD) and coefficient of variation (CV) for each of the parameters analysed using the calibration and validation sets after application of the CENTER algorithm and removal of any spectral outliers are shown in Table 1.

382 The nitrate content, ascorbic acid, and SSC reference values were all fairly
383 normally distributed around the mean values (1765.24 mg/kg, 298.97 mg/100 g and 8.84
384 ºBrix, respectively) with standard deviations (SD) of 1082.34 mg/kg, 66.05 mg/100g and
385 1.90 ºBrix for each parameter.
386 The ranges of nitrate content, ascorbic acid and SSC of the calibration set are
387 109.50-5177.00 mg/kg, 156.92-479.23 mg/100g and 5.60-14.25 ºBrix, respectively; that
388 of the validation set are 144.00-3520.00 mg/kg, 191.83-453.85 mg/100g and 6.25-13.95
389 ºBrix. Since the calibration and validation sets displayed similar values for mean, range
390 and standard deviation for all the parameters studied, a structured selection using only
391 spectral information treatment algorithms such as CENTER proved adequate.
392 Furthermore, the ranges of the validation set lay within those of the calibration set.
393 All the parameters studied in the calibration set covered a wide range of values.
394 This was truest of nitrate content with a CV of 63.31%, while the CV of ascorbic acid and
395 SSC were practically the same (22.09% and 21.49%, respectively). Pérez-Marín et al.
396 [43] have highlighted the importance of the sample set and of sample distribution within
397 the calibration set, noting that sample sets for calibration should ideally ensure an uniform
398 distribution across the range of the study parameter in question.
399 For the validation group, the coefficient of variation values were set at 60.62% for
400 nitrates, 21.10% for ascorbic acid and 22.66% for SSC.
401
402 *3.4. Prediction of safety and quality parameters in intact spinach using MPLS regression*
403
404 The best calibration models obtained for predicting one safety parameter (nitrate
405 content) and quality parameters (ascorbic acid and SSC) in spinach using different various

406 mathematical pre-treatments are shown in Table 2. Statistical criteria were used to select
407 the best model for each study parameter.

408 The models developed for nitrate content have a predictive capacity (r^2_{cv} = 0.41;
409 RDP_{cv} = 1.29) which allows the samples to be classified under high and low values of
410 this parameter [32, 39].

411 It should be noted that Xue and Yang [19] studied nitrate content in spinach (n =
412 58 samples), using an ASD Fieldspec FR spectroradiometer and obtained results (r^2_p =
413 0.88) which were higher than those obtained here, although the instrument's optical
414 characteristics and range are significantly different from those of the Phazir 2400.

415 Itoh et al. [20] also measured the nitrate content in spinach leaves (n = 48 samples),
416 using the FANTEC NIR Gun working on transmittance in a spectral range of 600-1100
417 nm. The authors reported values of RPD_p = 2.14 with the PCR regression and RPD_p =
418 2.17, using PLS regression, which were higher than those obtained in this research study.
419 However, the size and characteristics of the sample group, the means of measurement,
420 the window size (1 cm) and the spectral range of the instrument used, all differed from
421 those used in this study.

422 As regards ascorbic content, the models designed allow the samples to be
423 classified under high and low values for this parameter (r^2_{cv} = 0.33; RDP_{cv} = 1.21) [32,
424 39], although the results are limited for routine use. It must be noted that it is especially
425 difficult to measure this parameter in vegetable products - not only with portable
426 instruments, but also with high performance NIRS laboratory instruments.

427 Kramchote et al. [44] measured ascorbic acid content in cabbage and obtained
428 predictive capacity models (RPD_P = 1.26) similar to that obtained here using a
429 spectrophotometer Handy Lamda II (Spectra Co., Ltd., Tokyo, Japan), in reflectance
430 mode in the 310 and 1100 nm spectral range.

When predicting antioxidant content (beta-carotene and ascorbate) in freeze-dried leaves of *Populus* spp., Fernández-Martínez et al. [45] obtained models with an average predictive capacity ($RPD_{cv} = 2.10$) for beta-carotene, while for ascorbate, the models were capable of discriminating between high and low values ($RPD_{cv} = 1.56$) [26]. In both studies, reflectance was carried out using high performance Foss NIRS Systems 6500 laboratory equipment, in the 400-2500 nm range.

For SSC – which is a crucial parameter when choosing the optimum time for harvesting and for measuring the shelf-life of spinach, the model developed has a predictive capacity ($r^2_{cv} = 0.85$; $RPD_{cv} = 2.54$) which can be considered as good [32, 39].

No reports have been found in the literature regarding the measurement of SSC in spinach using NIR spectroscopy. However, Kramchote et al. [44], in the case of cabbage, obtained models of lower predictive capacity ($RPD_p = 2.05$ and 1.95) using a spectrophotometer (Handy Lamda II, Spectra Co., Ltd., Japan), in interactance and reflectance mode, respectively, in the spectral range 310-1100 nm.

The regression coefficients for the best predictive models for nitrate, ascorbic acid and soluble solids contents are illustrated in Fig. 2. The figure shows that the area of the spectra around 1650–1850 nm, which is correspond to the first overtone of C–H stretching bonds [42]. The best models to predict the three parameters tested reflected variations in the wavelengths range 2000–2190 to N–H and O–H stretching modes besides C=O vibration bands [42]. Absorbance region such as 2200-2280 nm and at 2320 nm could be attributed to C–H stretch and CH2 deformation [41].

Validations of the best calibration models obtained were performed using a set comprising 31 samples (Fig. 3). Models constructed for predicting SSC in intact spinach using the MEMS instrument met the validation requirements in terms of the coefficient of determination for prediction r^2_p ($r^2_p > 0.6$), while the standard error of prediction

456 corrected for bias or SEP(c), bias and slope were within the confidence limits: the
457 equation thus guarantees an accurate prediction, and can be applied routinely [40].

458 For nitrate content prediction, the r^2_p value does not comply with the protocol
459 while the values shown by the other statistics lie within the confidence limits, thus
460 complying with this validation protocol [40].

461 In the case of the ascorbic acid content, the r^2_p values and the slope do not comply
462 with this protocol, while SEP (c) and the bias were below confidence limits.

463

464 *3.5. Contrasting the suitability of performing reference analyses immediately after NIR*
465 *measurements*

466

467 Given the rough predictive capacity of the model designed to measure nitrate
468 content in spinach, and since the non-destructive prediction of this parameter is key in
469 determining the final destination of the harvested product, it was decided to evaluate
470 whether, for spinach leaves, carrying out chemical analyses at times other than when the
471 NIRS spectra are taken significantly affects the predictive capacity of the models. Several
472 authors have pointed out the importance of performing the reference analyses
473 immediately after the NIR analysis, especially in the case of extremely perishable
474 vegetable products [46].

475 Table 3 shows the best models obtained for predicting nitrates using two analysis
476 strategies: Strategy I: NIRS analysis, and chemical measurement of nitrate content 24
477 hours later; Strategy II: NIRS analysis and wet reference analysis immediately after.

478 Tables 2 and 3 clearly show the importance of carrying out the NIR and wet
479 process analyses consecutively, and of using the same methodology or systematic
480 analysis for all the samples in the study.

In this way, the joint predictive model (Table 2) shows values of r^2_{cv} = 0.41 and SECV = 836.26 mg/kg while the models developed from the two strategies considered (Table 3) had values of r^2_{cv} = 0.53 and 0.63, and SECV = 711.63 and 670.81 mg/kg, for strategies I and II, respectively.

It is important to point out that for both strategies, the SECV value was lower. In Strategy II, the SECV value fell by 19.78%, while for Strategy I, it decreased by 14.90%. Likewise, the predictive capacity of the models increased, and the models allowed to distinguish between high, medium and low values of nitrate content for both strategies [32, 39].

The results obtained indicate that growers and the agrifood industry could use NIRS technology as a screening technique, permitting a large number of plants to be tested in the field or when they are delivered to the industrial plant, providing results for this parameter and thus enabling growers to take real-time decisions as to the final destination of the harvested product.

These results also confirm the importance of performing these analyses together, in products as perishable as spinach [47] and the suitability of using the same analysis methodology throughout the trial.

3.6. Discriminant analysis

Results for the best classification models obtained, using PLS2-DA, for predicting the industrial use of spinach depending on its content in nitrates, are shown in Table 4.

The best discriminant models were obtained with D_2 log(1/R) together with SNV + DT for scatter correction (balanced and unbalanced sets).

505 The unbalanced model correctly classified 80.33% of the samples (85.71% in the
506 0-2,000 mg/kg category and 73.08% in the 2,000-3,500 mg/kg category), while the
507 balanced model correctly classified 75.00% (73.08% in the 0-2,000 mg/kg category and
508 76.92% in the 2,000-3,500 mg/kg category). For validation, the percentage of correctly
509 classified samples obtained using the unbalanced model was 70% (80% in the 0-2,000
510 mg/kg category and 60% in the 2,000-3,500 mg/kg category) and for the balanced model,
511 the percentage was 70.00% (90% in the 0-2,000 mg/kg category and 50% in the 2,000-
512 3,500 mg/kg category).

513 As can be seen in Table 4, the percentage of samples correctly classified by the
514 unbalanced model was slightly higher than by the balanced model, which reflects the low
515 sensitivity of the PLS2 algorithm to the difference between the number of samples of the
516 classes to be discriminated [48].

517 A detailed study of the unbalanced model reveals that, of the 12 poorly classified
518 samples, 5 had a nitrate content (mg/kg) of between 2,000 ± 2 * SEL (SEL = 131.6
519 mg/kg), a range that can be considered difficult to discriminate as it would fall within the
520 error of the reference method, so that the error of classification may be put down to a
521 typical laboratory error value. The remaining 7 samples did not show nitrate content
522 within this range, and therefore their incorrect classification can be attributed to errors in
523 the model or to their poor representation in the training set, given the limited number of
524 samples available, which only allowed to perform a feasibility study of the potential of
525 technology in this area. It is also important to mention that this constituent is found in
526 ppm in spinach leaves, which makes it difficult to measure by NIRS in whole plants.
527 However, the results obtained are promising and allow us to continue consolidating the
528 application of NIRS technology as a screening technique in the spinach handling and

processing industry, permitting, in a non-destructive way and in a matter of seconds, to assess the possible industrial destination of the spinach leaves.

Of the 13 poorly classified samples in the balanced model, 3 had a nitrate content (mg/kg) within the range of 2,000 ± 2 * SEL (SEL = 131.6 mg/kg), and any classification errors can again be attributed to the SEL value obtained.

Peaks at 1780 nm, 1860 nm and 2040 nm appeared to have more weight in the classification of intact spinach on-vine by nitrate content (Fig. 4). This indicates that the discrimination of spinach by nitrate content in the NIR region of the spectrum is related to lipids, proteins and N-H combinations [41].

As regards the external validation of the classification models, a percentage of correctly classified samples of 70% was obtained for both the unbalanced and balanced models. For the first, 10 samples from the 0-2,000 mg/kg category were used, of which 8 were correctly classified, and 10 from the 2,000-3,500 mg/kg category, of which 6 were correctly classified, while of the 4 poorly classified samples, 1 had a nitrate content (mg/kg) within the range 2,000 - 2 * SEL. Similarly, in the balanced model, 9 out of 10 samples in the 0-2,000 mg/kg category were correctly classified, and 5 from the 2,000-3,500 mg/kg category.

4. Conclusions

Near infrared spectroscopy is clearly an advantageous technique for the rapid screening of quality and safety according to the SSC and nitrate levels, although further research is needed to make it robust for predicting these parameters. It has also been demonstrated that the NIRS and the laboratory analysis should be performed together.

The results obtained from the classification models of spinach leaves according to their nitrate content, which determines their possible industrial destination, also confirm

555 the feasibility of using NIRS technology both in the field and in the stages of selection
556 and classification of spinach carried out during industrial processing for classification
557 according to the quality and safety characteristics of this vegetable.
558
559 **Acknowledgements**
560
561 This research was under the Research Project 'Quality determination of spinach
562 grown in Santaella (Córdoba)', funded by Gelagri Ibérica, S.L. The authors are grateful
563 to Mrs. María Carmen Fernández of the Animal Production Department for her technical
564 assistance. Furthermore, the authors wish to express their gratitude to the Spanish
565 Ministry of Education, Culture and Sports for the support offered to Irina Torres in the
566 form of the Training Programme for Academic Staff (FPU).
567
568 **References**
569 [1] G.H. Beis, A.S. Siomos, C.C. Dogras, Spinach composition as affected by leaf age
570 and plant part, Acta Hortic. 57 (2002) 653–658.
571 https://doi.org/10.17660/ActaHortic.2002.579.115.
572 [2] N. Smirnoff, Ascorbate biosynthesis and function in photo-protection, Phil. Trans.
573 355 (2000) 1455–1464. https://doi.org/10.1098/rstb.2000.0706.
574 [3] E. Koh, S. Charoenprasert, A.E. Mitchell, Effect of organic and conventional cropping
575 systems on ascorbic acid, vitamin c, flavonoids, nitrate, and oxalate in 27 varieties of
576 spinach (*Spinacia oleracea* L.), J. Agr. Food Chem. 60 (2012) 3144−3150.
577 https://doi.org/10.1021/jf300051f.
578 [4] G. Jaworska, Content of nitrates, nitrites, and oxalates in New Zealand spinach, Food
579 Chem. 89 (2005) 235−242. https://doi.org/10.1016/j.foodchem.2004.02.030.

580 [5] A. Elia, P. Santamaria, F. Serio, Nitrogen nutrition, yield and quality of spinach, J.
581 Sci. Food Agric. 76 (1998) 341–346. https://doi.org/10.1002/(SICI)1097-
582 0010(199803)76:3<341::AID-JSFA938>3.0.CO;2-4.
583 [6] P. Santamaria, Nitrate in vegetables: toxicity, content, intake and EC regulation, J.
584 Sci. Food Agric. 86 (2006) 10–17. https://doi.org/10.1002/jsfa.2351.
585 [7] J.M. Van Maanen, H.J. Albering, T.M. de Kok, S.G. van Breda, D.M. Curfs, I.T.
586 Vermeer, A.W. Ambergen, B.H. Wolffenbuttel, J.C. Kleinjans, H.M. Reeser, Does
587 the risk of childhood diabetes mellitus require revision of the guideline values for
588 nitrate in drinking water? Environ. Health Persp. 108 (2000) 457–461
589 https://doi.org/10.1289/ehp.00108457.
590 [8] J.L. Brown, N-Nitrosamines, J. Occup. Med. 14 (1999) 839–848.
591 [9] World Health Organization, Iron Deficiency Anaemia: Assessment, Prevention and
592 Control. A Guide for Programme Managers, WHO/NHD/01.3, World Health
593 Organization, Geneva,
594 http://apps.who.int/iris/bitstream/10665/66914/1/WHO_NHD_01.3.pdf?ua=1/,
595 2001 (accessed 3 June 2018).
596 [10] S. Bhuvaneswari, M. Joshi, A. D'Souza, Quantitative analysis of iron and ascorbic
597 acid contents in locally consumed fruits and vegetables, Int. Res. J. Biol. Sci. 4 (2015)
598 42–47.
599 [11] L. Hallberg, M. Brune, L. Rossander, The role of vitamin C in iron absorption, Int.
600 J. Vitam. Nutr. Res. 30 (1989) 103–108.
601 [12] Official Journal of the European Union (OJEU), Commission Regulation (EU) No
602 1258/2011 of 2 December 2011 Amending Regulation (EC) No 1881/2006 as regards
603 Maximum Levels for Nitrates in Foodstuffs, OJ L 320/15-17, 3.12.2011, 2011.

[13] S.K. Lee, A.A. Kader, Preharvest and postharvest factors influencing vitamin C content of horticultural crops, Postharvest Biol. Technol. 20 (2000) 207–220. https://doi.org/10.1016/S0925-5214(00)00133-2.

[14] A. Ito, H. Shimizu, H. Nakashima, J. Miyasaka, K. Ohdoi, Effect of different durations of root area chilling on the nutritional quality of spinach, IFAC Proc. Vol. 47 (2014) 4406–4410. https://doi.org/10.3182/20140824-6-ZA-1003.00744.

[15] H. Ito, H. Horie, K. Ippoushi, K. Azuma, Potential of visible-near infrared (VIS–NIR) spectroscopy for non-destructive estimation of nitrate content in Japanese radishes, Acta Hortic. 604 (2003) 549–552. https://doi.org/10.17660/ActaHortic.2003.604.64.

[16] H. Ito, F. Idezawa, Non-destructive determination of nitrate ion in leaf stalk of Qing gin cai using visible (VIS)-near infrared (NIR) spectroscopy, Acta Hortic. 712 (2006) 363–370. https://doi.org/10.17660/Acta Hortic.2006.712.41.

[17] S. Srivichien, A. Terdwongworakul, S. Teerachaichayut, Quantitative prediction of nitrate level in intact pineapple using Vis-NIRS, J. Food Eng. 150 (2015) 29–34. https://doi.org/10.1016/j.jfoodeng.2014.11.004.

[18] M.T. Sánchez, D. Pérez-Marín, I. Torres, B. Gil, A. Garrido-Varo, M.J. De la Haba, 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 (2017) 122–128. https://doi.org/10.1016/j.postharvbio.2016.11.011.

[19] L.H. Xue, L.Z. Yang, Nondestructive determination of nitrate content in spinach leaves with visible-near infrared high spectra, Spectrosc. Spect. Anal. 29 (2009) 926–930.

628 [20] H. Itoh, H. Tomita, Y. Uno, N. Shiraishi, Development of method for non-destructive
629 measurement of nitrate concentration in vegetable leaves by near-infrared
630 spectroscopy, IFAC Proc. Vol. 44 (2011) 1773–1778.
631 https://doi.org/10.3182/20110828-6-IT-1002.00738.
632 [21] A. Pissard, J.A. Fernández-Pierna, V. Baeten, G. Sinnaeve, G. Lognay, A. Mouteau,
633 P. Dupont, A, Rondia, M. Lateur, Non-destructive measurement of vitamin C, total
634 polyphenol and sugar content in apples using near-infrared spectroscopy, J. Sci. Food
635 Agric. 93 (2013) 238–244. https://doi.org/10.1002/jsfa.5779.
636 [22] M.T. Blanco-Díaz, M. Del Río-Celestino, D. Martínez-Valdivieso, R. Font, Use of
637 visible and near-infrared spectroscopy for predicting antioxidant compounds in
638 summer squash (*Cucurbita pepo* ssp *pepo*), Food Chem. 164 (2014) 301–308.
639 https://doi.org/10.1016/j.foodchem.2014.05.019
640 [23] C. Liu, S.X. Yang, L. Deng, Determination of internal qualities of Newhall navel
641 oranges based on NIR spectroscopy using machine learning, J. Food Eng. 161 (2015)
642 16–23. https://doi.org/10.1016/j.jfoodeng.2015.03.022
643 [24] R. Tierno, A. López, P. Riga, S. Arazuri, C. Jarén, L. Benedicto, J.I. Ruiz de
644 Galarreta, Phytochemicals determination and classification in purple and red fleshed
645 potato tubers by analytical methods and near infrared spectroscopy, J. Sci. Food
646 Agric. 96 (2016) 1888–1899. https://doi.org/10.1002/jsfa.7294.
647 [25] E.M. Toledo-Martín, M. García-García, R. Font, J.M. Moreno-Rojas, P. Gómez, M.
648 Salinas-Navarro, M. Del Río-Celestino, Application of visible/near-infrared
649 reflectance spectroscopy for predicting internal and external quality in pepper, J. Sci.
650 Food Agric. 96 (2016) 3114–3125. https://doi.org/10.1002/jsfa.7488.
651 [26] B.M. Nicolaï, K. Beullens, E. Bobelyn, A. Peirs, W. Saeys, K.I. Theron, J.
652 Lammertyn, Nondestructive measurement of fruit and vegetable quality by means of

653 NIR spectroscopy: a review, Postharvest Biol. Technol. 46 (2007) 99–118.
654 https://doi.org/10.1016/j.postharvbio.2007.06.024.
655 [27] M.T. Sánchez, D. Pérez-Marín, Nondestructive measurement of fruit quality by NIR
656 spectroscopy, in: M. Vázquez, J.A. Ramírez (Eds.), Advances in Post-Harvest
657 Treatments and Fruit Quality And Safety, Nova Science Publishers, Inc., Hauppauge,
658 NY, 2011, pp. 101–163.
659 [28] L.S. Magwaza, U.L. Opara, H. Nieuwoudt, P.J.R. Cronje, W. Saeys, B. Nicolaï, NIR
660 spectroscopy applications for internal and external quality analysis of citrus fruit–a
661 review, Food Bioprocess Tech. 5 (2012) 425–444. https://doi.org/10.1007/s11947-
662 011-0697-1.
663 [29] R.B. Thompson, M. Gallardo, M. Joya, C. Segovia, C. Martínez-Gaitán, M.R.
664 Granados, Evaluation of rapid analysis systems for on-farm nitrate analysis in
665 vegetable cropping, Span. J. Agric. Res. 7 (2009) 200–211.
666 https://doi.org/10.5424/sjar/2009071-412.
667 [30] K. Tsukazawa, Study of simplicity measurement methods of deoxidize type vitamin
668 C and B-carotene content in vegetables, Bull. Saitama Pref. Agr. For. Res. Ctr. 2
669 (2002) 43–46.
670 [31] J.S. Shenk, M.O.Westerhaus, Analysis of Agriculture and Food Products by Near
671 Infrared Reflectance Spectroscopy, Monograph, NIRSystems, Inc., 12101 Tech
672 Road, Silver Spring, MD 20904, USA, 1995.
673 [32] J.S. Shenk, M.O.Westerhaus, Calibration the ISI way, in: A.M.C. Davies, P.C.
674 Williams (Eds.), Near Infrared Spectroscopy: The Future Waves, NIR Publications,
675 Chichester, 1996, pp. 198–202.

676 [33] M.L. Martínez, A. Garrido, E.J. De Pedro, L. Sánchez, Effect of sample
677 heterogeneity on NIR meat analysis: the use of the RMS statistic, J. Near Infrared
678 Spectrosc. 6 (1998) 313–320. https://doi.org/10.1255/jnirs.214.
679 [34] M. Rosales, Use of the NIR Spectral Information per se for Quality Evaluation of
680 Agricultural Products, Master Thesis, University of Cordoba, Cordoba, 1993.
681 [35] R.J. Barnes, M.S. Dhanoa, S.J. Lister, Standard normal variate transformation and
682 de-trending of near infrared diffuse reflectance spectra, Appl. Spectrosc. 43 (1989)
683 772–777. https://doi.org/10.1366/0003702894202201.
684 [36] J.S. Shenk, M.O. Westerhaus, Routine Operation, Calibration, Development and
685 Network System Management Manual, NIRSystem, Inc., 12101 Tech Road, Silver
686 Spring, MD 20904, USA, 1995.
687 [37] ISI, The Complete Software Solution Using a Single Screen for Routine Analysis,
688 Robust Calibrations, and Networking, Manual, FOSS NIRSystems/TECATOR,
689 Infrasoft International, LLC, Sylver Spring MD, USA, 2000.
690 [38] J.S. Shenk, M.O. Westerhaus, Population structuring of near infrared spectra and
691 modified partial least squares regression, Crop Sci. 31 (1991) 1548–1555.
692 [39] P.C. Williams, Implementation of near-infrared technology, in: P.C. Williams, K.H.
693 Norris (Eds.), Near-Infrared Technology in the Agricultural and Food Industries,
694 AACC Inc., St. Paul, 2001, pp. 145–169.
695 [40] W.R. Windham, D.R. Mertens, F.E. Barton II, Protocol for NIRS calibration: sample
696 selection and equation development and validation, in: G.C. Martens, J.S. Shenk,
697 F.E. Barton II (Eds.), Near Infrared Spectroscopy (NIRS): Analysis of Forage
698 Quality, Agriculture Handbook, vol. 643, USDA-ARS, US Government Printing
699 Office, Washington, DC, 1989, pp. 96–103.

700 [41] J.S. Shenk, J.J.Jr. Workman, M. Westerhaus, Application of NIR spectroscopy to
701 agricultural products, in: D.A. Burns, E. Ciurczak (Eds.), Handbook of Near-Infrared
702 Analysis, Marcel Dekker, Basel, 2008, pp. 347–386.
703 [42] B.G. Osborne, T. Fearn, P.H. Hindle, Practical NlR spectroscopy with applications,
704 in: B.G. Osborne, T. Fearn, P.H. Hindle (Eds.), Food and Beverage Analysis,
705 Longman, Essex, 1993, pp. 11–35.
706 [43] D. Pérez-Marín, A. Garrido-Varo, J.E. Guerrero, Implementation of LOCAL
707 algorithm with near-infrared spectroscopy for compliance assurance in compound
708 feedingstuffs, Appl. Spectrosc. 59 (2005) 69–77.
709 https://doi.org/10.1366/0003702052940585.
710 [44] S. Kramchote, K. Nakano, S. Kanlayanarat, S. Ohashi, K.Takizawa, G. Bai, Rapid
711 determination of cabbage quality using visible and near-infrared spectroscopy, LWT
712 Food Sci. Technol. 59 (2014) 695–700. https://doi.org/10.1016/j.lwt.2014.07.009.
713 [45] J. Fernández-Martínez, R. Joffre, M. Zacchini, B. Fernández-Marín, J.I. García-
714 Plazaola, I. Fleck, Near-infrared reflectance spectroscopy allows rapid and
715 simultaneous evaluation of chloroplast pigments and antioxidants, carbon isotope
716 discrimination and nitrogen content in *Populus* spp. leaves, Forest Ecol. Manag. 399
717 (2017) 227–234. https://doi.org/10.1016/j.foreco.2017.05.041.
718 [46] D.C. Pérez-Marín, A. Garrido, J.E. Guerrero, A. Gómez-Cabrera, Near-infrared
719 reflectance spectroscopy (NIRS) for the mandatory labelling of compound
720 feedingstuffs: chemical composition and open-declaration, Anim. Feed Sci. Technol.
721 116 (2004) 333–349. https://doi.org/10.1016/j.anifeedsci.2004.05.002.
722 [47] M. Cantwell, T. Suslow, Spinach: Recommendations for Maintaining Postharvest
723 Quality,

http://postharvest.ucdavis.edu/Commodity_Resources/Fact_Sheets/Datastores/Vegetables_English/?uid=32&ds=799, 2002 (accessed 5 June 2018).

[48] D. Pérez-Marín, A. Garrido-Varo, J.E. Guerrero, Optimization of discriminant partial least squares regression models for the detection of animal by-product meals in compound feeding stuffs by near-infrared spectroscopy, Appl. Spectrosc. 60 (2006) 1432-1437. https://doi.org/10.1366/000370206779321427.

Table 1

Minimum, maximum and mean values for the RMS and STD statistics and the STD limit value obtained while evaluating spectral repeatability.

Number of samples	RMS (µlog(1/R)			STD (µlog(1/R)			STD_{limit} (µlog(1/R)
	Minimum	Mean	Maximum	Minimum	Mean	Maximum	
10	4,971	34,147	86,210	7,030	48,292	121,919	64,661
20	34,626	83,157	139,285	47,334	118,693	214,102	128,437

737 **Table 1**

738 Statistical analysis of the calibration and prediction sample sets, i.e. data ranges, means

739 and standard deviations (SD) and coefficients of variation (CV).

Parameter	Item	Calibration set	Validation set
Nitrate content (mg/kg)	Number	93	31
	Range	109.50-5177.00	144.00-3520.00
	Mean	1765.24	1795.48
	SD	1082.34	1088.46
	CV (%)	61.31	60.62
Ascorbic acid content (mg/100 g)	Number	93	31
	Range	156.92-479.23	191.83-453.85
	Mean	298.97	300.96
	SD	66.05	63.52
	CV (%)	22.09	21.10
SSC (º Brix)	Number	93	31
	Range	5.60-14.25	6.25-13.95
	Mean	8.84	8.34
	SD	1.90	1.89
	CV (%)	21.49	22.66

740

741 **Table 2**

742 Calibration statistics for the best equations obtained for the prediction of quality and

743 safety parameters in intact spinach.

Parameter	Number of samples	Mathematical treatment	Range	Mean	SD	r^2_{cv}	SECV	RPD	SEL
Nitrate (mg/kg)	92	2,5,5,1	109.50-5177.00	1776.16	1083.10	0.41	836.26	1.29	131.61
Ascorbic acid (mg/100 g)	91	2,10,5,1	168.75-467.69	298.55	62.24	0.33	51.46	1.21	10.69
SSC (ºBrix)	92	1,10,5,1	5.60-14.25	8.85	1.91	0.85	0.75	2.54	0.07

744

745

746 **Table 3**

747 Calibration statistics for the best equations obtained for the prediction of nitrate content

748 in intact spinach. Strategies I and II.

Strategy	Number of samples	Mathematical treatment	Range	Mean	SD	r^2_{cv}	SECV	RPD_{cv}
Strategy I	44	2,5,5,1	109.50-5177.00	1648.82	1037.97	0.53	711.63	1.46
Strategy II	71	1,5,5,1	122.50-3627.00	1821.91	1095.37	0.63	670.81	1.63

749

750

751 **Table 4**

752 Percentage of spinach plants classified correctly by nitrogen content. PLS2-DA.

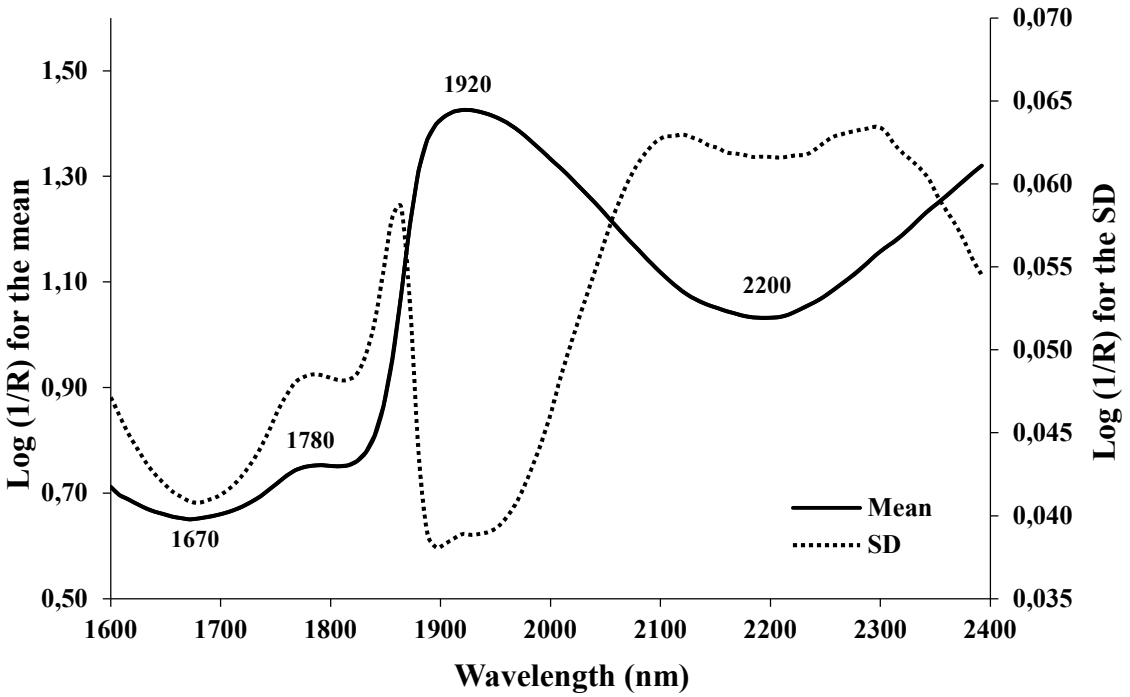
Qualitative groups	**Unbalanced model**		**Balanced model**	
	Percentage of correctly-classified samples: 80.33% (49/61) Model SECV: 0.46 Number of synthetic variables: 4 Math treatment: 2,5,5,1		Percentage of correctly-classified samples 75.00% (39/52) Model SECV: 0.49 Number of synthetic variables: 5 Math treatment: 2,10,5,1	
Industrial destination of spinach according to its content in nitrates	Training set	Validation set	Training set	Validation set
Preserved, deep-frozen or frozen spinach, NO_3^- : 0-2,000 mg /kg	85.71% (30/35)	80.00% (8/10)	73.08% (19/26)	90.00% (9/10)
Fresh spinach NO_3^- : 2,000-3,500 mg /kg	73.08% (19/26)	60.00% (6/10)	76.92% (20/26)	50.00% (5/10)

753

754 **Fig. 1.** Mean and standard deviation spectrum for spinach.

755

756



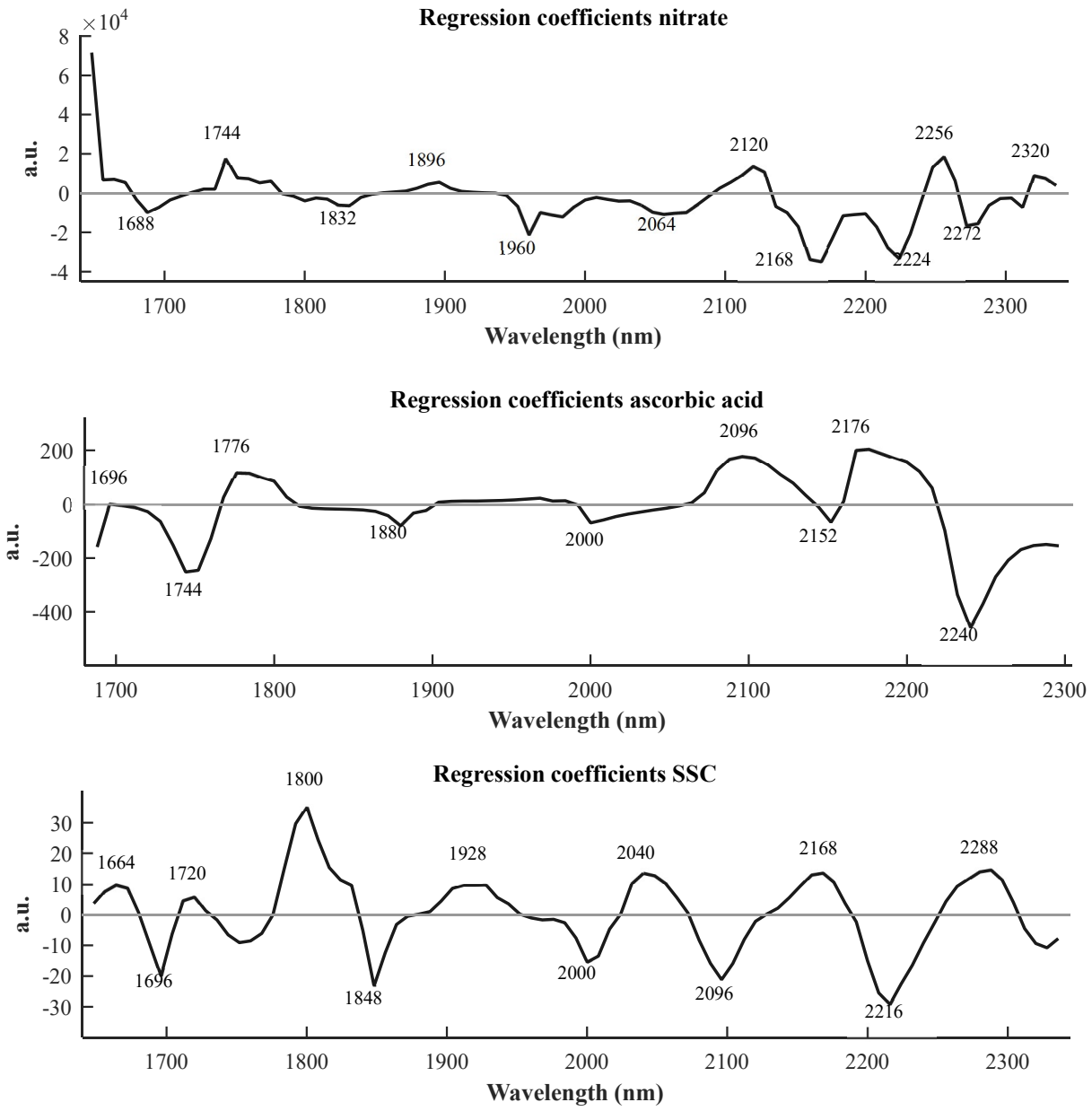
757

758

759 **Fig. 2.** Regression coefficients for spinach nitrate content, ascorbic acid and SSC during

760 on-vine ripening

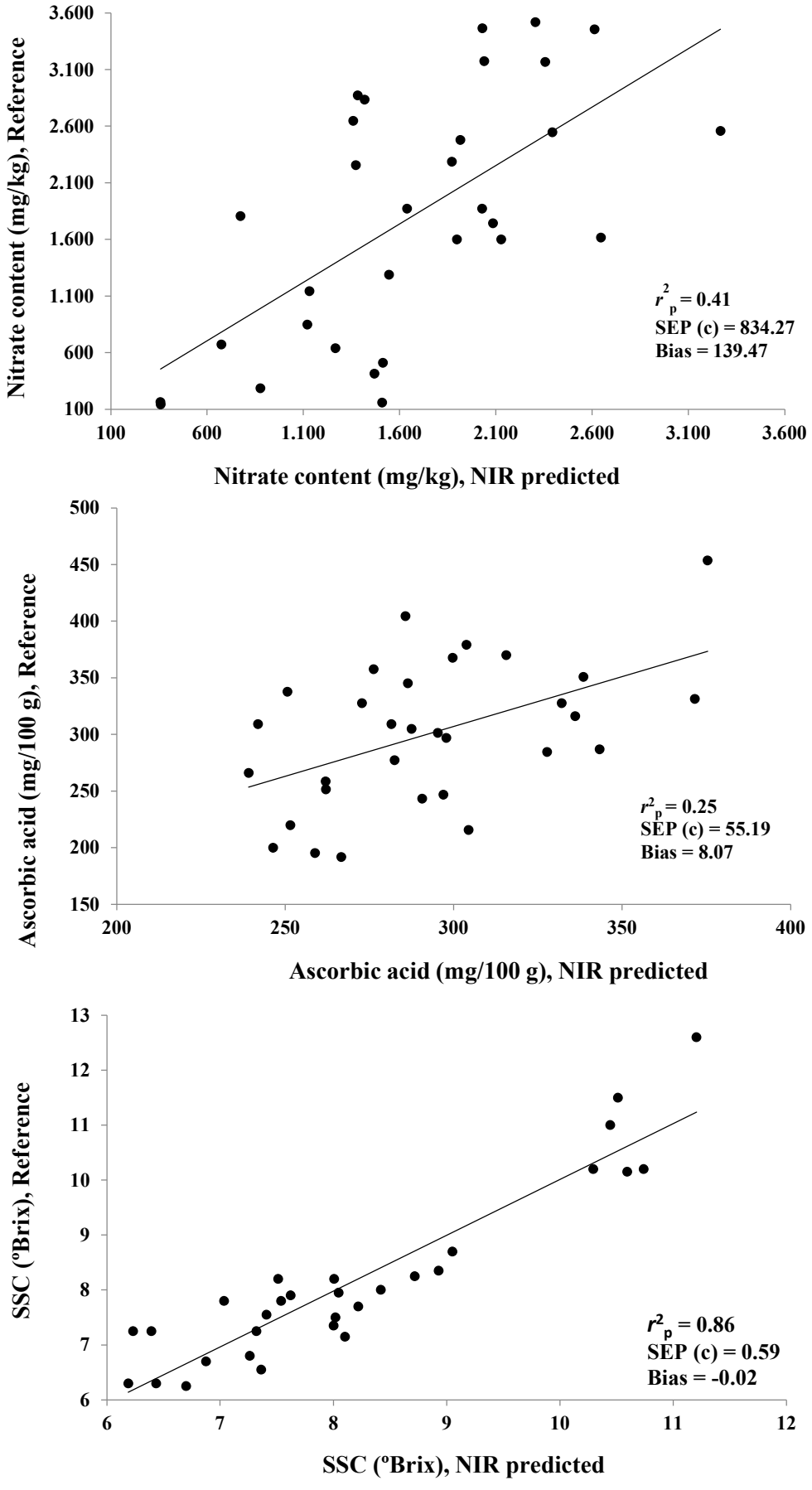
761



762 * a.u.= arbitrary units

763

764 **Fig. 3.** Reference and NIR predicted values for quality and safety parameters in intact
765 spinach.



766

767 **Fig. 4**. D_1 log(1/R) of the mean spectra for intact spinach plants with different nitrate

768 content.

769