

*Estimation of total soluble solids in grape berries using a hand held
NIR spectrometer under field conditions*

AUTHOR'S NAMES

Ruben Urraca^a, Andres Sanz-Garcia^b, Javier Tardaguila^c and Maria P Diago^{c*}

a Universidad de La Rioja, 26006, Logroño, Spain.

b Division of Biosciences. University of Helsinki. Viikinkaari, 5 E (P.O. Box 56), 00014, Helsinki, Finland.

c Instituto de Ciencias de la Vid y del Vino (University of La Rioja, CSIC, Gobierno de La Rioja), 26006 Logroño, Spain

*Correspondence to: Maria P Diago. Instituto de Ciencias de la Vid y del Vino (University of La Rioja, CSIC, Gobierno de La Rioja), 26006 Logroño, Spain

E-mail: mpaz.diago.santamaria@gmail.com

Telephone: +34941299760

Fax:+34941299721

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/jsfa.7470

ABSTRACT

BACKGROUND

Recent studies have reported the potential of near infrared (NIR) spectral analyzers for monitoring the ripeness of grape berries alternatively to wet chemistry methods. This study covers various aspects regarding the calibration and implementation of predictive models of total soluble solids (TSS) in grape berries using laboratory and in-field collected NIR spectra.

RESULTS

The performance of the calibration models obtained under laboratory conditions indicated that at least 700 berry samples are required to assure enough prediction accuracy. A statistically significant error reduction ($\Delta\text{RMSECV}=0.1$ °Brix) with $p<0.001$ was observed when measuring berries without epicuticular wax, which was negligible from a practical point of view. Under field conditions, the prediction errors ($\text{RMSEP}=1.68$ °Brix, $\text{SEP}=1.67$ °Brix) were close to those obtained with the laboratory dataset ($\text{RMSEP}=1.42$ °Brix, $\text{SEP}=1.40$ °Brix).

CONCLUSION

This work clarifies several methodological factors to develop a protocol for in-field assessing TSS in grape berries using an affordable, non-invasive, portable NIR spectral analyzer.

KEYWORDS LIST

NIR spectroscopy; precision viticulture; non invasive sensor; chemometrics; total soluble solids content

INTRODUCTION

Monitoring the ripening of grape berries at several timings during three to four weeks prior to harvest is a current practice in the wine industry worldwide. Among the ripeness variables, total soluble solids (TSS) accumulation in the berries is a prerequisite for the subsequent alcohol content after fermentation in the wine, but in-field sampling and subsequent laboratory analysis of the must obtained from the crushed berries is needed. Although widely adopted in the grape

and wine industry, two main concerns accompany this method. The first and most critical one is the representativeness of the berry sample (usually 100 to 200 berries, or 15-25 whole clusters) from a given vineyard plot. Secondly, the destructive and time consuming nature of the analysis (conducted in the laboratory), which may prevent it from being replicated as many times as needed along the season. To overcome these issues, the use of non-invasive sensors in the field, to assess grape composition parameters is a matter of great interest in recent years for the wine industry.

Near infrared (NIR) spectroscopy is a powerful, non-invasive technique that is being increasingly applied in food industry thanks to the development of cheaper, faster and more accurate sensors.¹ It has come to the market as a rapid and affordable technology to replace the manual classification or other tedious wet chemistry analyses. Some of the NIR applications already implemented in industry include on-line quality control systems^{2,3} or multi-function sensors to monitor product properties at once⁴.

In the wine industry, so far, several works on the application of NIR spectroscopy to wine and grapes have already been carried out with promising results. In grape musts, different parameters were measured: total soluble solids (TSS),⁵ phenolic compounds,⁶⁻⁸ pH,^{5,9} or titratable acidity¹⁰ among others. Initial studies assessed maturity with spectra collected from must after crushing and mixing different grapes during the ripening process.^{11,12} Although accurate estimations were obtained, the potential of NIR technology had not been fully exploited, as sample preparation was still required. The presentation mode of the berry samples was simplified in subsequent works by collecting spectra directly on individual picked grapes without crushing the berries^{10,13-22} under controlled laboratory conditions. The ultimate step was to acquire NIR spectra in vineyards, directly on-the-vine, using portable equipment. Very few exploratory studies have addressed this approach,²³⁻²⁵ significantly hindered by the varying conditions of field measurements. In the work of Larraín et al.²⁴ the validation models showed R^2 varying from 0.87 to 0.93 with size ranges of 144 and 740 samples in Carmenere and Chardonnay cultivars, when a portable spectrometer operating between 700 to 1100 nm was used. However, no discussion on the suitable number of samples or the TSS range was reported. A different device, working in the range of 1600 to 2400 nm, was used by González-Caballero et al.²³, but due to the limited number of samples, the leave-one-out cross-validation

was used, and no results from the obtained models were reported. Neither the influence of naturally-occurring substances attached to the berry skin, like the epicuticular wax, has been fully addressed. Nevertheless some test were undertaken by Larraín et al.²⁴ to measure the influence of the dust on the quality of the measurements. In the line of the work conducted by Mukhtar et al.²⁶ with plums, a more exhaustive research is still required to characterize the epicuticular wax influence in NIR-based predictions.

From a practical application of portable NIR devices under field conditions, to monitor berry ripening from veraison to harvest, several factors such as the minimum number of berry samples to build the model, the range of the parameter to be assessed, the influence of the epicuticular wax, have not been fully addressed

Besides, as far as the equipment used for spectra collection is concerned, most of the cited studies are based on the use of non-portable laboratory equipment such as diode-array^{10, 17, 18} or monochromators,^{8, 9} which hinders the extrapolation of this methods under field conditions. When used to scan samples directly on-the-vine, some authors adapted the laboratory devices to the field or they even manufactured their own prototypes of portable NIR spectrometers.^{15, 24, 25} All these devices did not hold the required properties for a day-to-day ripening monitoring process. On the other hand, new portable NIR devices based on micro-electro-mechanical systems (MEMS) technology are nowadays being manufactured at lower cost to ease their implementation in industrial applications. However, as reported by Sánchez et al.²⁷, limited information is available on the agricultural and food sector in regard to NIR-MEMS technology. Only few authors have addressed the use of MEMS technology in real case studies with several types of fruit^{3, 27, 28} in general, and with grapes^{13, 23} in particular. There is still a need to clarify all the aforementioned aspects regarding the calibration models using this type of portable devices, and to develop a practical and handy methodology to boost the in-field implementation of portable, low-cost NIR spectroscopy based on MEMS by wineries.

The present work provides a comprehensive comparison of the different possibilities and methodologies to use these portable NIR sensors working in the range of 1600 to 2500 nm under field conditions to non-invasively monitor the ripening status of grape berries by means of the TSS to meet industrial applications. The goal is to evaluate the influence of various methodological factors and to define several practical and methodological guidelines to

implement NIR technology under field conditions by analyzing the following aspects: (1) number of samples required to calibrate the model, (2) influence of the TSS range of the grape berries in the calibration process, (3) impact of the spectral pre-processing techniques in the accuracy of predictions, (4) influence of the epicuticular wax on spectra acquisition, and (5) selection of an adequate procedure to implement the technology under field conditions without picking berries.

MATERIALS AND METHODS

Samples

A database of NIR (1600-2400 nm) reflectance data obtained from collected spectra and total soluble solids content of Tempranillo (*Vitis vinifera* L.) berries was created under laboratory conditions. Sixty-eight clusters were manually harvested between September and October 2013 at different locations of Rioja wine appellation in Spain. For each cluster, 25 berries were randomly picked for spectra collection, resulting in a final database of 1600 samples. All spectra were acquired with a NIR portable spectral analyzer (microPHAZIR™, Thermo Fisher Scientific Inc., Waltham, MA). Afterwards, the same berries were individually hand-crushed for total soluble solids (TSS, expressed in °Brix) measurement, which was performed using a digital refractometer (model WM-7, Atago CO., LTD, Tokyo, Japan).

A subset of 100 additional berries was created to study the influence of the epicuticular wax covering the berry skin²⁶ on spectra collection. For this purpose, grape berries with virtually intact epicuticular wax were selected. Two spectra were collected for each berry. The first one was obtained on parts of the berry with intact epicuticular wax. The surface of the whole berry was then cleaned with a kimwipe, and the second spectrum was acquired. To maintain similar conditions in spectra acquisition, the second shot was roughly taken in the same parts of the berry as those used for the first one.

For the in-field dataset, Tempranillo (*Vitis vinifera* L.) berries were measured in the field with the NIR portable spectral analyzer (microPHAZIR, Thermo Fisher Scientific Inc., Waltham, MA, USA) while they were still on the hanging clusters of the grapevines of a commercial vineyard in Logroño (La Rioja, 42°26'51"N, 2°30'06"O), Spain. This database of 43 samples was built up

with these in field measurements to externally validate the models calibrated with the laboratory database. All berries were measured within a single day throughout different vines of the field. After spectra acquisition, each berry was promptly placed inside a numbered plastic bag and transported to the laboratory under refrigerated conditions (10°C). Prior to TSS measurement, berries were allowed to stabilize at room temperature.

Spectra collection

An integrated handheld NIR spectral analyzer (microPHAZIR™, Thermo Fisher Scientific Inc., Waltham, MA), designed to analyze diffuse reflection, was used for spectra collection in reflectance mode ($\log 1/R$) both in the laboratory and in the field. The analyzer covers the wavelength range from 1595.7 to 2396.3 nm in intervals of 8.7 nm (100 bands). Sensor integration time was 600 ms. The device was equipped with quartz protection to prevent dirt accumulation. Three spectra were obtained per sample and the average spectrum was used thereafter. Samples of laboratory database were presented as whole berries. Each individual berry extracted from the cluster was placed in contact to the analyzer reflection window (Fig. 1a). The window was cleaned between measurements to avoid contamination between samples. Under field conditions (Fig. 1b), the same presentation mode was replicated but berries were not detached from the hanging cluster on the vines for spectrum acquisition. Berries were measured when the reflection window of the NIR device was fully touching one specific berry of the cluster.

All spectra were retrieved from the device, saved in ASCII format and imported to R software²⁹ for the subsequent analysis.

Model calibration under laboratory conditions

The statistical R software was used for the chemometric analysis with packages hyperSpec,³⁰ proSpectr,³¹ pls³² and vegan.³³ Before calibration, data were analyzed through principal component analysis (PCA).³⁴ In order to improve outlier detection with PCA, a multiple scatter correction (MSC)^{35, 36} filter was used prior to data projection. Outliers were identified based on the 95 % confidence ellipse.³⁷ In a further step, additional samples were discarded by optimizing the predicting models based on the residuals.^{1, 38}

After outlier detection, the influence of the following factors on the quality of predictions was evaluated:

- *Size of the calibration set.* the predictive models were trained with 18 different calibration sets, comprised by a variable number of samples. Spectra were mean centered and the calibration set size varied from 50 to 900 samples in intervals of 50 samples.
- *Influence of TSS range.* A benchmark composed by 13 calibration sets was created. The TSS range in the different calibration sets varied from 10.3 to 21.6 °Brix. Calibration sets were obtained by performing a proportional stratified sampling between a minimum and a maximum value of TSS. The maximum was kept constant and equal to the highest TSS value in the laboratory database (29 °Brix) while the minimum was shifted from 4 to 16 °Brix in 1 °Brix intervals. Data were mean centered and the number of samples was set to 700.
- *Spectral pre-processing techniques.*^{39, 40} The following spectral pre-processing techniques were compared; a) mean centering (MC); b) multiple scatter correction (MSC); c) standard normal variate (SNV);⁴¹ d) de-trending (DT) using first, second and third order polynomials,⁴¹ and e) the zero first and second order Savitzky-Golay (SG) filter using 7 points of smoothing and a second order polynomial.^{42, 43} A calibration set of 700 samples was used.
- *Influence of the epicuticular wax.* Spectra from intact and cleaned berries were mean centered and tested with the calibrated models. Results obtained were statistically compared using the parametric paired t-test.⁴⁴ The Shapiro-Wilk test was prior used to test the normality of distributions.⁴⁵

In all situations, partial least squares (PLS) regression was the predictive technique chosen.^{46, 47} The whole spectral range was used as inputs for all calibrated models. Besides, in the majority of cases and for the calibration of the final predictive model, the pre-processing technique mean centering and a calibration set of 700 samples were used according to the results obtained after analyzing these factors separately. The prediction errors of models were estimated by k-fold Cross-Validation (CV).⁴⁸ The final k value used was 10 in order to balance the bias and computational efficiency on the basis of the size of the calibration set. The number

of latent variables (LV) was sequentially increased until reaching the first relative minimum in the root mean squared error of cross validation (RMSECV). Hence, models with a minimum number of LV were developed avoiding over-fitting.⁴⁹ An external validation set was used to test model prediction capabilities. The size of the external validation set was the 30 % of the corresponding calibration set used. To assess the effect of the presence of the epicuticular wax on the berries, external validation was not implemented due to the low number of available samples. Instead, the 10 CV procedure was repeated 100 times (100 x 10 CV) to perform a proper statistical comparison.⁵⁰ For each predictive model, the statistics computed were: coefficient of determination (R^2), root mean square error (RMSE), standard error (SE) and bias. All of them were obtained for calibration (C), cross validation (CV) and prediction (P) sets. Besides, the residual predictive deviation (RPD),⁵¹ which is defined as the ratio between the standard deviation of the prediction set (SD) and the standard error of prediction (SEP), was calculated. A RPD higher than 3 is widely associated to a good predictive ability of calibration models for screening processes.⁵²

External validation against field data

Outlier detection in the field database was performed using PCA. First, the 95% confidence ellipse was computed and displayed over the PC score plot of laboratory samples. Then, field samples were added to the PC score plot using the PC loadings obtained with the laboratory samples. Field samples that fell out of the ellipse were automatically discarded. The best calibration model trained using laboratory data was tested with field data and the statistics listed above for the calibration process were again computed.

RESULTS

Figure 2 shows the NIR spectra of both, laboratory (Fig. 2.a) and field (Fig. 2.b) acquisitions. All field spectra fell within the interval of laboratory spectra, enabling the implementation of a model calibrated with the laboratory spectra under field conditions. A strong peak was observed around 1950 nm due to OH absorption, reflecting mainly water.⁸ Looking at the interquartile range (shadow area), a constant variation was observed in the range scanned, shrinking only in the region before the OH peak. As shown in Table 1, TSS values in the general laboratory set

ranged from 4.6 to 29.9 °Brix, as sample picking extended from incipient veraison to harvest. The TSS ranged from 9.8 to 26.7 °Brix in the subset of 100 samples selected for analyzing the influence of the epicuticular wax. Based on the 95% confidence interval, 30 samples of the general set and 4 of the epicuticular wax subset were removed. Additionally, 10% of the remaining samples were discarded based on the residuals plot. This procedure led to a general set of 1323 samples and an epicuticular wax subset of 85 samples. Calibration and prediction sets for each experiment (Table 1) were created by proportional stratified sampling from the original cleaned general set throughout the whole TSS range.

Size of the calibration set

Figure 3.a shows the evolution of R^2 and RMSE in the predictive models trained with 18 different calibration sets of different number of samples. A significant variation of both, R^2 and RMSE, was observed for a number of samples, n , lower than 250. Calibration and CV errors stabilized when the size of the calibration set, n , surpassed 250 samples. However, RMSEP and R^2P still exhibited substantial oscillations for calibration sets between 250 and 700 samples ($250 < n < 700$). As expected, they both improved as the size increased; however, prediction errors did not stabilize until the size of calibration set reached 700 samples ($n > 700$).

Range of the dependent variable in the calibration database

Figure 3.b depicts the evolution of the absolute and relative errors for different ranges of the dependent variable in the calibration set, that is, RMSE and R^2 respectively. Thirteen different calibration sets were used with different values of TSS range: 10.3, 11.3, 12.3, 13.2, 14.3, 15.3, 16.3, 17.3, 18.2, 19.3, 20.2, 21 and 21.6 °Brix. Results showed that all RMSEs slightly improved when models were trained with lower values of TSS range. RMSEP went from 1.51 to 1.31 °Brix for calibration ranges from 21.6 to 10.3 °Brix. However, the relative error (R^2) substantially declined when the TSS range in the calibration set diminished. The value of R^2P dropped from 0.90 to 0.73 for calibration ranges of 21.6 and 10.3 °Brix, respectively.

Pre-processing techniques

Eight pre-processing techniques were compared using the whole spectrum and a calibration set of 700 samples. The option of no pre-processing the spectra was also included in the study and labeled as raw. Models were optimized based on their internal validation results, but the

performance of each pre-processing technique was evaluated using the external validation set. In all situations, the training errors (C) were close to the expected generalization errors (CV) and real generalization errors (P). This proved an absence of over-fitting during the training process. As shown in Table 2, the techniques with the lowest prediction errors were SG0 using 13 latent variables (RMSEP=1.46 °Brix, SEP=1.52 °Brix), DT1 using 12 latent variables (RMSEP=1.47 °Brix, SEP=1.53 °Brix), MC using 16 latent variables (RMSEP=1.42 °Brix, SEP=1.40 °Brix) and the option of no pre-processing, with 15 latent variables (RMSEP=1.45 °Brix, SEP=1.52 °Brix). Moreover, most of the calibrations showed a RPD > 3 which confirmed the good performance of the models.

MC exhibited the best predictive capacity with a RPD = 3.36 with 16 latent variables, which accounted for the 99 % of variance in X and 95 % of variance in Y. The number of latent variables was selected by optimizing the RMSECV. PLS loading weights of the four first latent variables of this model are shown in Figure 4.b. The first LV remained relatively constant throughout the whole range of wavelengths, which revealed the importance of the integral component as a predictor. However, the next loading showed a significant variation within the range of 1800 - 2100 nm, where the absorption peak due to water occurred. The regression coefficients depicted in Figure 4.a also remarked the importance of the range highlighted before.

Influence of the epicuticular wax on spectra acquisition

Boxplots in Figure 5 depict the RMSECV and R^2CV distribution after performing a 100 x 10 CV with the subset for epicuticular wax analysis. The RMSECV for intact berries ranged from 1.32 to 1.47 °Brix with an average of 1.40° Brix, whereas this range was extended from 1.22 to 1.43 °Brix with a mean RMSECV of 1.31 °Brix after cleaning the berries. The differences between means of RMSECVs and R^2CV s in models trained with cleaned and intact berries were evaluated using the parametric paired t-test. Prior the use of the t-test, data normality was checked by Saphiro-Wilk test (p-value > 0.05). The test yielded p-values for cleaned and intact berries of 0.498 and 0.487 for both RMSECV distributions and 0.309 and 0.567 for R^2CV , enabling the use of the paired t-test. Statistically significant differences were found in both parameters between the intact and cleaned berries datasets (p-value < 0.001).

Model implementation under field conditions

TSS values for grape berries in field data ranged from 15 to 23.5 °Brix, which was within the range of the laboratory set used for calibration (Table 1). This field set was used to externally validate the best model calibrated with laboratory samples, i.e. a PLS model trained with mean centered data and calibration set of 700 samples. For outlier detection, field samples were plotted over the principal components (PC) score plot of the laboratory data used to train the predictive model (Fig. 6). No outliers were identified with this procedure as all field points fell within the 95% confidence ellipse of laboratory data.

Predictions obtained with the laboratory model in field data were plotted against the real field values in Figure 7. These results were compared against the predictions obtained with same model in the laboratory prediction (P) set. Under field conditions, a RMSEP=1.68 °Brix and a SEP=1.67 °Brix were obtained; close to the errors obtained with the external validation set from the laboratory (RMSEP=1.42 °Brix and SEP=1.40 °Brix). However, the R^2P under field conditions ($R^2P=0.38$) was considerably lower than the one obtained in the laboratory ($R^2P=0.91$). This was a consequence of the influence of the distribution of data in the coefficient of determination, as this statistic evaluates the quantity of variance (SD^2) explained by the model. This variance was substantially higher in the laboratory prediction set ($SD^2=22.03$ °Brix²) than in the field set ($SD^2=4.61$ °Brix²) because field samples were collected within a single day. Therefore, and in order to facilitate an even comparison of the results, all statistics were again computed in a subset of the laboratory prediction set where similar conditions to those of the field prediction set were simulated. This new subset was sampled from the laboratory prediction set ($n=175$) with the same size ($n=43$) and a similar range and variance ($SD^2=4.21$ °Brix²) than the field set. In this sampled subset, the RMSEP remained roughly constant (RMSEP=14.8 °Brix) but the R^2P drastically decreased ($R^2P=0.47$), as it occurred with the field prediction set.

DISCUSSION

The viability of monitoring the grape ripening status using NIR technology without the necessity of crushing grapes prior to measurements has already been proved by several authors.^{10, 13, 14, 18, 19, 24} However, more research on how to translate these results to in-situ vineyard monitoring was needed to evaluate the potential of NIR portable systems. For this

purpose, the present work was designed to investigate into several methodological factors to apply NIR technology under field conditions for industrial application, rather than outperforming the existing studies in terms of accuracy. Hence, a PLS model was trained with a calibration set built under laboratory conditions, with the aim of having higher quality spectra, and subsequently tested with in-field collected spectra, directly on-the-vine. In terms of accuracy, the errors obtained with our calibration set ($R^2CV=0.89$, $RMSECV=1.51$ °Brix, $R^2P = 0.91$; $RMSEP = 1.41$ °Brix) were similar to those reported by Guidetti et al.²⁵ ($RMSEP = 1.48$ °Brix), Bellincontro et al.¹³ ($RMSEP = 0.72$ °Brix; $R^2P = 0.92$), Barnaba et al.¹⁹ ($RMSEP = 0.93$ °Brix, $R^2P = 0.94$), González-Caballero et al.¹⁰ ($R^2CV = 0.91$) and Cao et al.¹⁴ ($RMSEP = 0.93$ °Brix; $R^2P = 0.91$). In these studies, spectra were directly collected from intact berries under controlled laboratory conditions but with more costly and complex analyzers, operating in different wavelength ranges, mainly designed for indoor spectra acquisition.

Among the methodological factors to be studied, the size of the calibration dataset was the first one attempted, as the number of samples required to calibrate the model was a key parameter. Similar works, where the intact berries were used, displayed some discrepancies regarding this factor; from the 251 samples used by González-Caballero et al.¹⁸, through 108 used by González-Caballero et al.¹⁰, 96 used by Barnaba et al.¹⁹, 450 used by Bellincontro et al.¹³ to the 3,135 samples used by Kemps et al.¹⁷. Although it might seem evident *a priori* that the larger the number of samples the more accurate and stable predictions could be obtained, the effort required to create these calibration sets must also be taken into account. Our results showed that a calibration set with less than 250 samples was clearly insufficient. Acceptable results in terms of stability were obtained with calibration sets between 250 and 700 samples, and no further improvement was appreciated beyond 700 samples. Consequently, a minimum of 250 samples is required to perform any reliable exploratory analysis, while 700 or more samples would be sufficient to develop a robust model for practical applications of the model as an on-line sensor.

Another important factor when building the calibration set was the TSS range of the grape berries sampled. In this regard, the value of R^2 significantly increased along the TSS range. For a virtually constant absolute error (RMSE), the relative error (R^2) varied due to the change on variance caused by the different values of the TSS range. These results suggest that

comparison between models trained with different datasets, even if the same spectral range is covered, must be cautiously done. In order to ensure a fair comparison, the same calibration conditions must be followed, i.e, same variety, similar number of samples and same TSS range. Therefore, results of the present study, which included Tempranillo berries throughout the whole industrial ripening process up to 29.9 °Brix, cannot be properly compared with other studies in which even raisins have been sampled.^{10, 18}

The potential noise introduced by the epicuticular wax in the spectra was the last factor discussed regarding calibration. Statistically significant lower prediction errors were found after cleaning the berries' surface ($p < 0.001$), but this improvement was overlooked under a normal use of TSS values. The increase in accuracy in terms of RMSE was 0.1 °Brix, which does not compensate for the effort put in cleaning the berries. Moreover, the removal of this protective layer accelerates the water loss of the berry and worsens its visual appearance.²⁶

Consequently, authors agree with the preliminary trials conducted by Larraín et al.²⁴, who reported that no statistical difference was observed between sampling grape berries with or without dust, which ultimately had a similar effect as the epicuticular wax in spectra acquisition. The lack of necessity of berry cleaning prior to NIR spectra acquisition with the portable device is an important outcome that can facilitate the application of this non-invasive technology under field conditions.

Once the optimally sized calibration set was determined and other methodological aspects clarified, the next step was to perform a chemometric analysis. After outlier removal, different pre-processing techniques were evaluated. Previous studies have utilized a wide variety of these techniques and even combinations of them. Several authors reported that the best predictions were obtained with SNV plus detrending,¹⁰ first order derivative,^{17, 20} mean centering,¹³ mean centering plus SNV¹⁹ or even no pre-processing.^{14, 24} According to our results, the best performing techniques were Savitzky-Golay filter of zero order, detrending using a first order polynomial, mean centering and the option of no pre-processing (raw). Following an approach based on the parsimony principle,⁵³ in which the simplest method should be selected when several are equivalent in terms on accuracy and precision, no pre-processing could have been selected *a priori*. However, MC was the chosen option because more stable predictions were obtained with PLS regression when data were MC and presented as variations

around the mean.¹ This increase on stability was clearly appreciated with a bias reduction in the prediction set from -0.22 to 0.04 when MC was used. In fact MC spectra is almost similar to the not pre-processed data, as MC is just a requirement when using PLS. Hence, our findings agree with the results obtained by Bellincontro et al.¹³, Larraín et al.²⁴, who reported that not significant improvement was observed using various pre-processing techniques. In this way, Nicolai et al.³⁸ also stated that there did not seem to be a large advantage of either transformation method compared to the untransformed data, and the acquisition of spectra of upmost quality was stressed.

The best calibration model with the laboratory database was used to predict the TSS contents in grape berries using their spectra collected directly on-the-vine under field conditions. A small difference in terms of RMSEP and SEP between laboratory and field errors proved the suitability of NIR spectroscopy to perform in-field measurements of TSS directly on-the-vine. The low R^2P obtained was a direct consequence of collecting all field samples within a single day, which led to low variance ($SD^2=4.61$ °Brix²) compared to the one of the real population ($SD^2=22.03$ °Brix²). This fact was demonstrated by simulating the existing conditions in the field set in a subset sampled from the laboratory prediction set. In this new subset, R^2P drastically dropped to 0.47 despite the initial value of 0.91 computed using the whole laboratory prediction set. Conversely, RMSEP did not experience a significant variation (RMSEP = 1.42 vs. RMSEP = 1.48 °Brix). This proved that low R^2 values, such as those obtained using field data, do not necessarily imply a bad fit of the model. It is important to note that R^2 is only a valid statistic when a dataset with range and variance close to the true variance and range of the real population studied is available. Besides, and in line with the results of the TSS range analysis, all of this suggests that absolute errors such as RMSE can provide a more appropriate level of assertiveness than R^2 when evaluating models performance through different datasets.⁴⁸ Consequently, the RMSE is preferable than R^2 when evaluating the quality of predictions obtained in a short period of time.

Our results demonstrate the suitability of portable NIR spectroscopy operating in the range of 1600 to 2500 nm, to be used under field conditions to assess the TSS in grape berries. Additionally, this type of portable NIR sensors may be used to characterize the spatial variability of sugar accumulation at any time prior to harvest if used in combination with a GPS system,

which could allow the georeferencing of the measurements. This information can also be of great importance in the frame of the application of precision viticulture strategies aimed at performing selective harvesting to increase the profitability and quality of the wines. Overall, the proficiency of NIR portable spectroscopy under field conditions to assess the TSS in the berries was proved and practical and methodological aspects to ensure a proper use of this technology were studied.

CONCLUSIONS

Results obtained in the present study entail an additional step for the use of portable NIR spectral analyzers as non-destructive sensors to monitor the total soluble solid content in grape berries under field conditions. A deeper understanding of various methodological aspects for the assessment of TSS using non-invasive portable NIR devices was provided. In this regard, a minimum of 700 intact berry samples, covering a TSS range of 18 °Brix was recommended to build a robust calibration spectral set. Of the different spectral pre-processing techniques, the simplest options, which were mean centering and no-pre-process, were those showing the best accuracy performance. The presented results are of special interest for the grape and wine industry as the use of this non-invasive technology enables the monitorization of sugar accumulation across the ripening process at many points as needed to ensure their representativeness at no expense of yield reduction and time consumption.

ACKNOWLEDGEMENTS

This work has received funding from the European Union's Seventh Programme for research, technological development and demonstration under grant agreement N 311775 (KBBE-INNOVINE).

One of the authors (ASG) would like to acknowledge the support provided by the "Instituto de Estudios Riojanos (IER)", the Research Funding No. 273689 (FINSKIN) and the mobility grant No. 276371 (VATURP) from the Academy of Finland. Special thanks to Juan Fernández-Novales for his careful revision of the manuscript.

Accepted Article

REFERENCES

1. Cozzolino D, Cynkar WU, Shah N and Smith P, Multivariate data analysis applied to spectroscopy: Potential application to juice and fruit quality. *Food Res Int* **44**:1888-1896 (2011).
2. Nansen C, Zhang X, Aryamanesh N and Yan G, Use of variogram analysis to classify field peas with and without internal defects caused by weevil infestation. *J Food Eng* **123**:17-22 (2014).
3. Pérez-Marín D, Paz P, Guerrero J-E, Garrido-Varo A and Sánchez M-T, Miniature handheld NIR sensor for the on-site non-destructive assessment of post-harvest quality and refrigerated storage behaviour in plums. *J Food Eng* **99**:294-302 (2010).
4. Pan L, Zhu Q, Lu R and McGrath JM, Detection of sucrose content of sugar beet by visible/near infrared spectroscopy, in *Proceedings of the American Society of Agricultural and Biological Engineers Annual International Meeting*, Ed, Kansas City, MO, pp 1-9 (2013).
5. Damberg RG, Cozzolino D, Cynkar WU, Janik L and Gishen M, The determination of red grape quality parameters using the LOCAL algorithm. *J Near Infrared Spec* **14**:71-79 (2006).
6. Ferrer-Gallego R, Hernández-Hierro JM, Rivas-Gonzalo JC and Escribano-Bailón MT, Determination of phenolic compounds of grape skins during ripening by NIR spectroscopy. *LWT-Food Sci Technol* **44**:847-853 (2011).
7. Frago S, Aceña L, Guasch J, Busto O and Mestres M, Application of FT-MIR spectroscopy for fast control of red grape phenolic ripening. *J Agr Food Chem* **59**:2175-2183 (2011).
8. Janik L, Cozzolino D, Damberg RG, Cynkar WU and Gishen M, The prediction of total anthocyanin concentration in red-grape homogenates using visible-near-infrared spectroscopy and artificial neural networks. *Anal Chim Acta* **594**:107-118 (2007).
9. Cozzolino D, Cynkar WU, Damberg RG, Janik L and Gishen M, Effect of both homogenisation and storage on the spectra of red grapes and on the measurement of total anthocyanins, total soluble solids and pH by visual near infrared spectroscopy. *J Near Infrared Spec* **13**:213-223 (2005).
10. González-Caballero V, Sánchez MT, López MI and Pérez-Marín D, First steps towards the development of a non-destructive technique for the quality control of wine grapes during on-vine ripening and on arrival at the winery. *J Food Eng* **101**:158-165 (2010).
11. Fernández-Navales J, López M-I, Sánchez M-T, García-Mesa J-A and González-Caballero V, Assessment of quality parameters in grapes during ripening using a miniature fiber-optic near-infrared spectrometer. *Int J Food Sci Nutr* **60**, **Supplement 7**:265-277 (2009).
12. Picque D, Lieben P, Chretien P, Beguin J and Guerin I, Assessment of maturity of loire valley wine grapes by mid-infrared spectroscopy. *J Int Sci Vigne Vin* **44**:219-229 (2010).
13. Bellincontro A, Cozzolino D and Mencarelli F, Application of NIR-AOTF spectroscopy to monitor aleatico grape dehydration for passito wine production. *Am J Enol Viticult* **62**:256-260 (2011).
14. Cao F, Wu D and He Y, Soluble solids content and pH prediction and varieties discrimination of grapes based on visible-near infrared spectroscopy. *Comput Electron Agric* **71**, **Supplement 1**:S15-S18 (2010).
15. Herrera J, Guesalaga AR and Agosin E, Shortwave-near infrared spectroscopy for non-destructive determination of maturity of wine grapes. *Meas Sci Technol* **14**:689-697 (2003).
16. Jarén C, Ortuño JC, Arazuri S, Arana JI and Salvadores MC, Sugar determination in grapes using NIR technology. *Int J Infrared Milli* **22**:1521-1530 (2001).
17. Kemps B, Leon L, Best S, De Baerdemaeker J and De Ketelaere B, Assessment of the quality parameters in grapes using VIS/NIR spectroscopy. *Biosyst Eng* **105**:507-513 (2010).
18. González-Caballero V, Pérez-Marín D, López MI and Sánchez MT, Optimization of NIR spectral data management for quality control of grape bunches during on-vine ripening. *Sensors* **11**:6109-6124 (2011).
19. Barnaba FE, Bellincontro A and Mencarelli F, Portable NIR-AOTF spectroscopy combined with winery FTIR spectroscopy for an easy, rapid, in-field monitoring of Sangiovese grape quality. *J Sci Food Agr* **94**:1071-1077 (2014).
20. Arana JI and Arazuri S, Maturity, variety and origin determination in white grapes (*Vitis Vinifera* L.) using near infrared reflectance technology. *J Near Infrared Spec* **13**:349-357 (2005).
21. Damberg RG, Cozzolino D, Esler MB, Cynkar WU, Kambouris A, Francis IL, Høj PB and Gishen M, The use of near infrared spectroscopy for grape quality measurement. *Austr New Zealand Grapegr Wine* **473a**:69-76 (2003).
22. Casiraghi E, Sinelli N, Bodria L, Guidetti R, Beghi R and Cabassi G, Evaluation of grape quality parameters by Vis/NIRs and FT-NIRs spectroscopy. *Ital Food Bev Technol* **50**:5-10 (2007).

23. González-Caballero V, Sánchez MT, Fernández-Navales J, López MI and Pérez-Marín D, On-vine monitoring of grape ripening using near-infrared spectroscopy. *Food Anal Method* **5**:1377-1385 (2012).
24. Larraín M, Guesalaga AR and Agosín E, A multipurpose portable instrument for determining ripeness in wine grapes using NIR spectroscopy. *IEEE T Instrum Meas* **57**:294-302 (2008).
25. Guidetti R, Beghi R and Bodria L, Evaluation of grape quality parameters by a simple Vis/NIR system. *Trans American Soc Agric Biol Engin* **53**:477-484 (2010).
26. Mukhtar A, Damerow L and Blanke M, Non-invasive assessment of glossiness and polishing of the wax bloom of European plum. *Postharvest Biol Tec* **87**:144-151 (2014).
27. Sánchez MT, De la Haba M-J and Pérez-Marín D, Internal and external quality assessment of mandarins on-tree and at harvest using a portable NIR spectrophotometer. *Comput Electron Agric* **92**:66-74 (2013).
28. Sánchez MT, De la Haba M-J, Benítez-López M, Fernández-Navales 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).
29. R-Core-Team, R: A language and environment for statistical computing, Ed. R foundation for statistical computing, Vienna, Austria (2014).
30. Beleites C and Sergo V, hyperSpec: a package to handle hyperspectral data sets in R. *J Stat Softw* (2012).
31. Stevens A and Ramirez-Lopez L, prospectr: miscellaneous functions for the processing and sample selection of vis-NIR spectral data, Ed (2013).
32. Mevik BH and Wehrens R, The pls package: Principal component and partial least squares regression in R. *J Stat Softw* **18**:1-24 (2007).
33. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Gavin L, Solymos P, Stevens MHM and Wagner H, vegan: community ecology package, Ed (2013).
34. Wehrens R, *Chemometrics with R. Multivariate data analysis in the natural sciences and life sciences*. Springer-Verlag, New York, NY (2011).
35. Martens H, Jensen SA and Geladi P, Multivariate linearity transformations for near infrared reflectance spectroscopy, in *Proceedings of the Nordic Symposium on Applied Statistics*, ed. by Christie OHH, Stokland Forlag, Stavanger, Norway, pp 205-234 (1983).
36. Geladi P, MacDougall D and Martens H, Linearization and scatter-correction for Near-Infrared Reflectance spectra of meat. *Appl Spectrosc* **39**:491-500 (1985).
37. Eriksson L, Byrne T, Johansson E, Trygg J and Vikstrom C, *Multi- and megavariate data analysis basic principles and applications*. Umetrics, Malmo, Sweden (2013).
38. Nicolai 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 Tec* **46**:99-118 (2007).
39. Rinnan Å, van den Berg F and Engelsen SB, Review of the most common pre-processing techniques for near-infrared spectra. *Trends Anal Chem* **28**:1201-1222 (2009).
40. Rinnan Å, Nørgaard L, Berg Fvd, Thygesen J, Bro R and Engelsen SB, Chapter 2 - Data Pre-processing, in *Infrared Spectroscopy for Food Quality Analysis and Control*, ed. by Sun D-W. Academic Press, San Diego, CA, pp 29-50 (2009).
41. 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).
42. Savitzky A and Golay MJE, Smoothing and differentiation of data by simplified least squares procedures. *Anal Chem* **36**:1627-1639 (1964).
43. Steiner J, Termonia Y and Deltour J, Smoothing and differentiation of data by simplified square procedure. *Anal Chem* **44**:1906-1909 (1972).
44. Crawley MJ, *The R Book*. John Willey & Sons, Ltd, Chichester, West Sussex, UK (2013).
45. Shapiro SS and Wilk MB, An analysis of variance test for normality (complete samples). *Biometrika* **52**:591-611 (1965).
46. Geladi P and Kowalski BR, Partial least-squares regression: A tutorial. *Anal Chim Acta* **185**:1-17 (1986).
47. Wold S, Sjöström M and Eriksson L, PLS-regression: a basic tool of chemometrics. *Chemometr Intell Lab* **58**:109-130 (2001).
48. Kuhn M and Johnson K, *Applied predictive modelling*. Springer, New York, NY (2013).
49. Martens H and Naes T, *Multivariate calibration*. John Willey & Sons, Ltd, Chichester, West Sussex, UK (1992).
50. Rafaelizadeh P, Tang L and Liu H, Cross validation, in *Encyclopedia of Database Systems*, ed. by Liu L and Ozsu MT. Springer, New York, NY, pp 1-6 (2009).

51. Williams PC, Implementation of near-infrared technology, in *Near-Infrared technology in the Agricultural and Food Industries*, ed. by Williams PC and Norris KH. AACC Inc, St. Paul, MN, pp 125-269 (2001).
52. Williams PC and Sobering DC, How do we do it: a brief summary of the methods we use in developing near infrared calibrations, in *Near Infrared Spectroscopy: The future waves, the proceedings of the 7th International Conference on Near Infrared Spectroscopy*, Ed by Davies AMC and Williams P. NIR Publications, Montreal, pp 158-188 (1996).
53. Hastie T, Tibshirani R and Friedman J, *The elements of statistical learning: Data mining, inference and prediction*. Springer, New York, NY (2009).

FIGURES' CAPTIONS



Figure 1. Spectra acquisition in (A) laboratory, and field conditions (B) using the portable microPHAZIR™ (ThermoScientific, Waltham, MA, USA) NIR spectral analyzer, used for non-invasive assessment of total soluble solids.

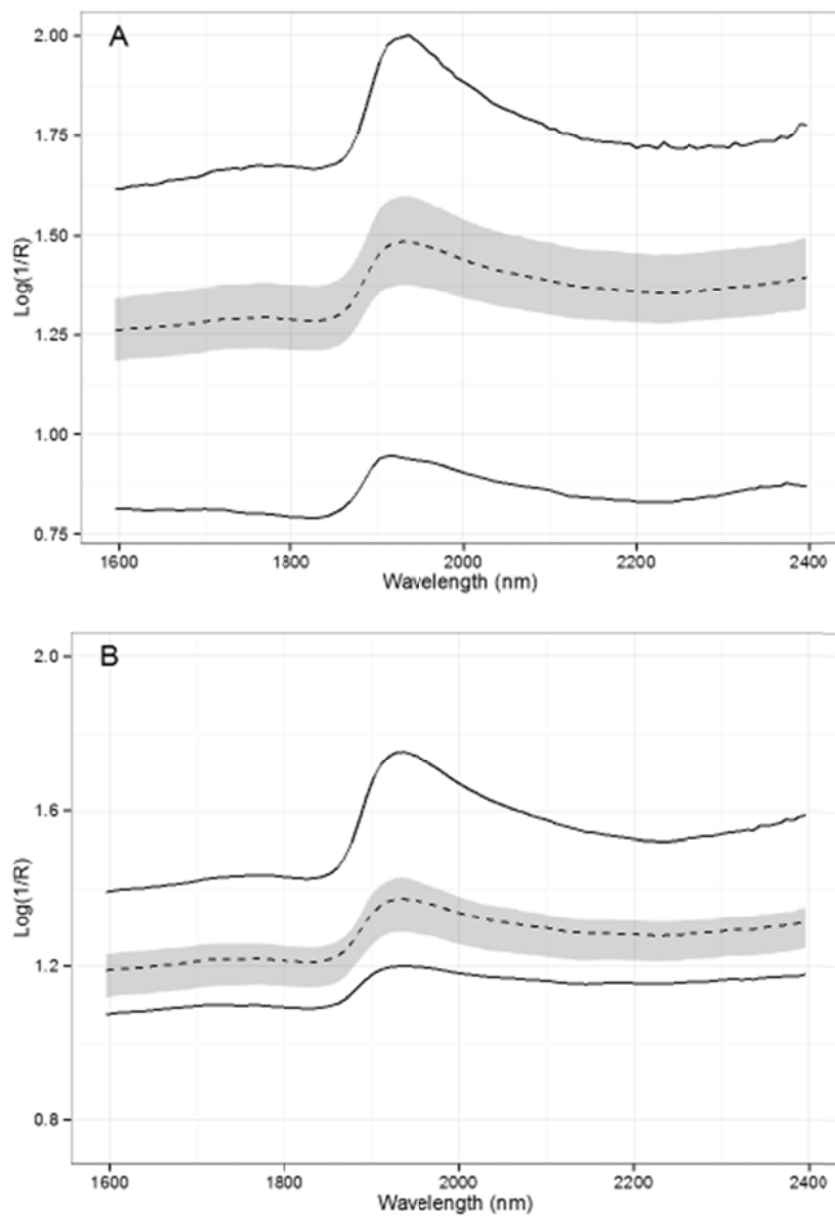


Figure 2. Mean, minimum, maximum and interquartile range (shadow area) of absorbance spectra of berries scanned under laboratory (A) and in-field (B) conditions.

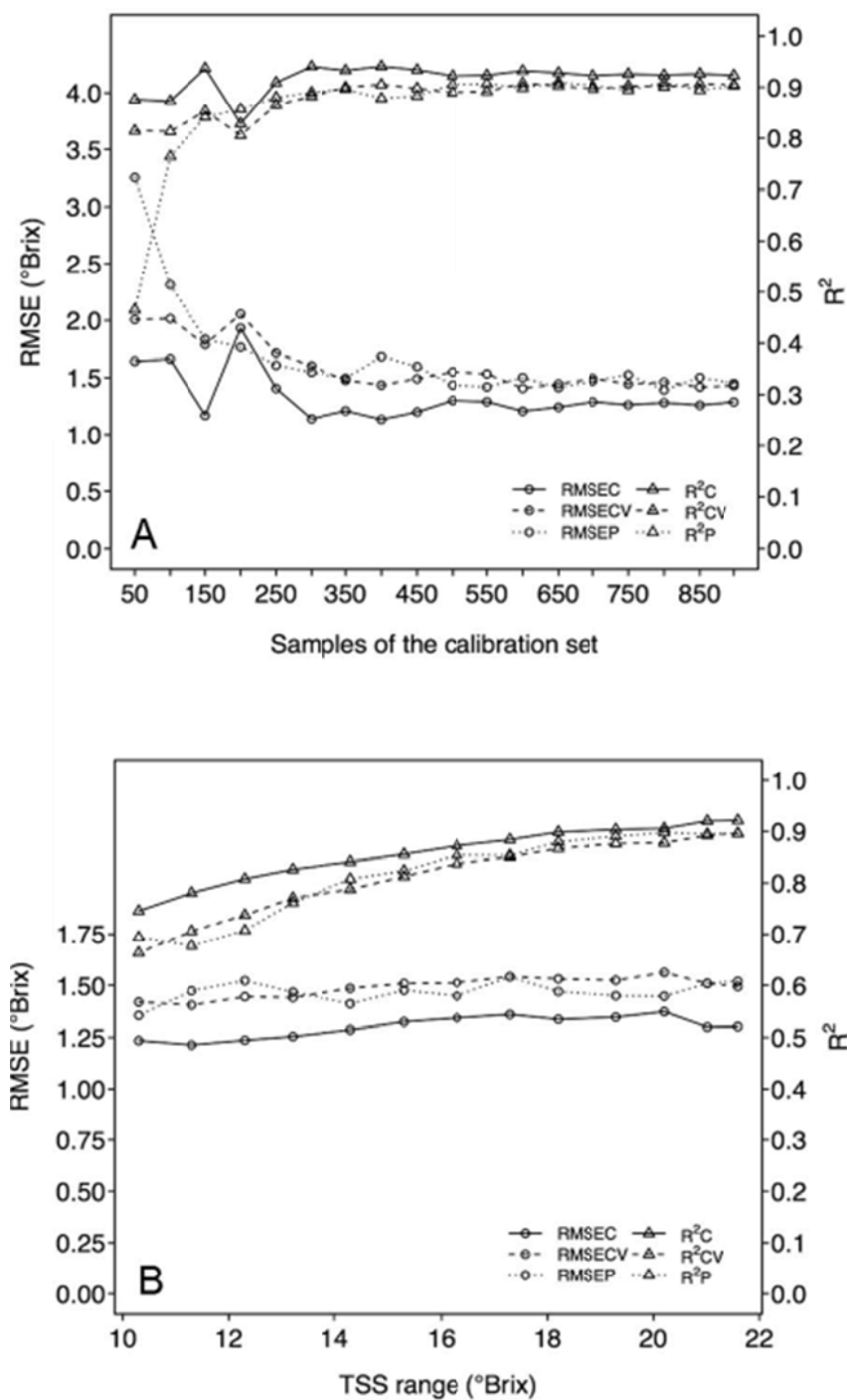


Figure 3. RMSE and R² values for the calibration (RMSEC and R²C), cross validation (RMSECV and R²CV) and prediction (RMSEP and R²P) models versus the number of

samples (A) and versus de TSS range (B) used in the calibration set for training the PLS models.

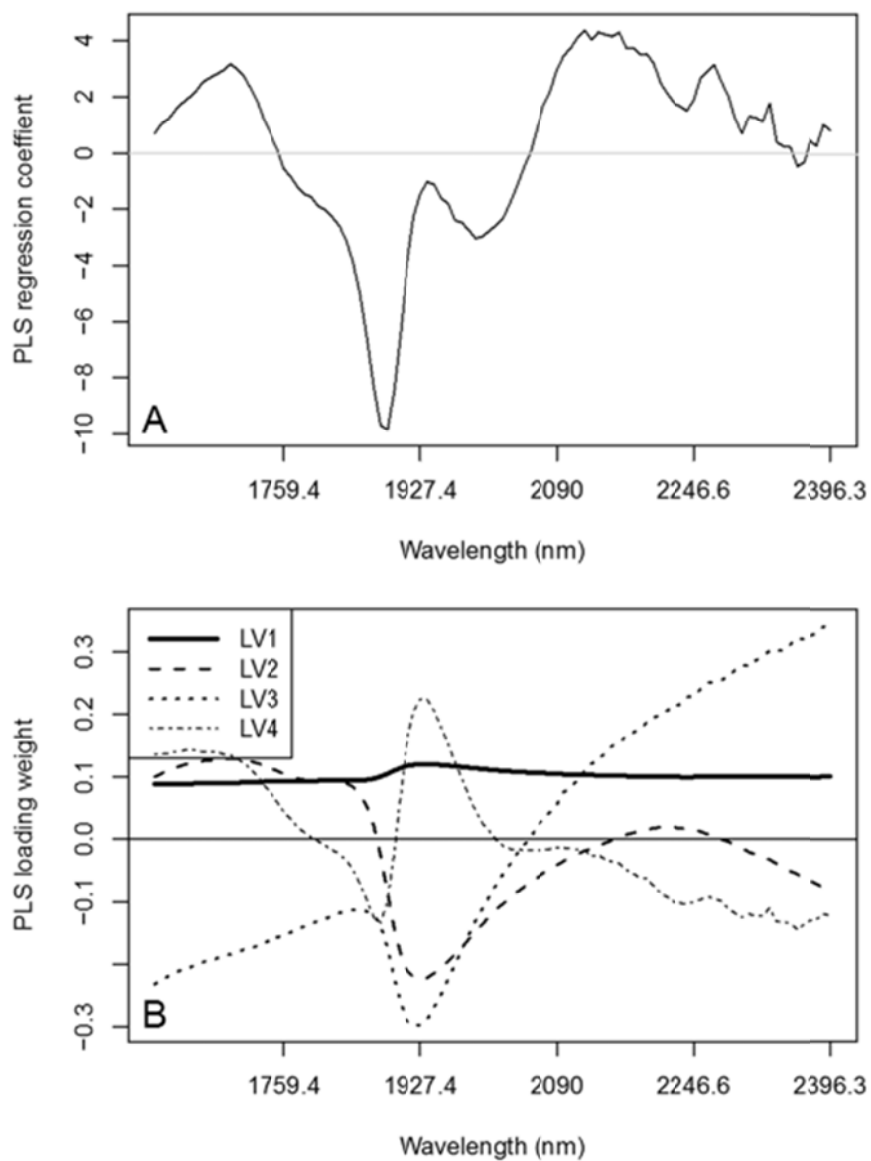


Figure 4. Regression coefficients (A), and loading weights (B) of the RMSECV with the number of latent variables of the best PLS laboratory model.

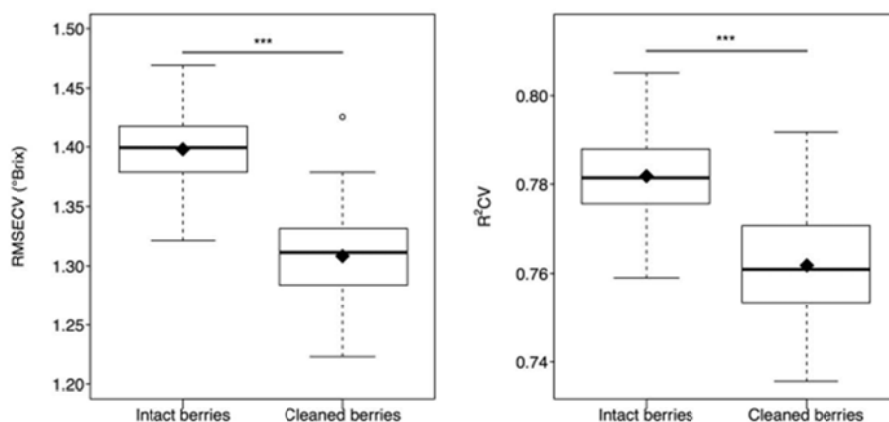


Figure 5. Boxplots representing the distribution of RMSECV and R^2CV of models created using the intact and cleaned (after removal of the epicuticular wax) berries databases with $n=95$ components after 100×10 CV ($t = -20.278$, $p\text{-value} = 2.2 \times 10^{-16}$ for RMSECVs, and $t = -13.2313$, $p\text{-value} = 2.2 \times 10^{-14}$ for R^2CV s).

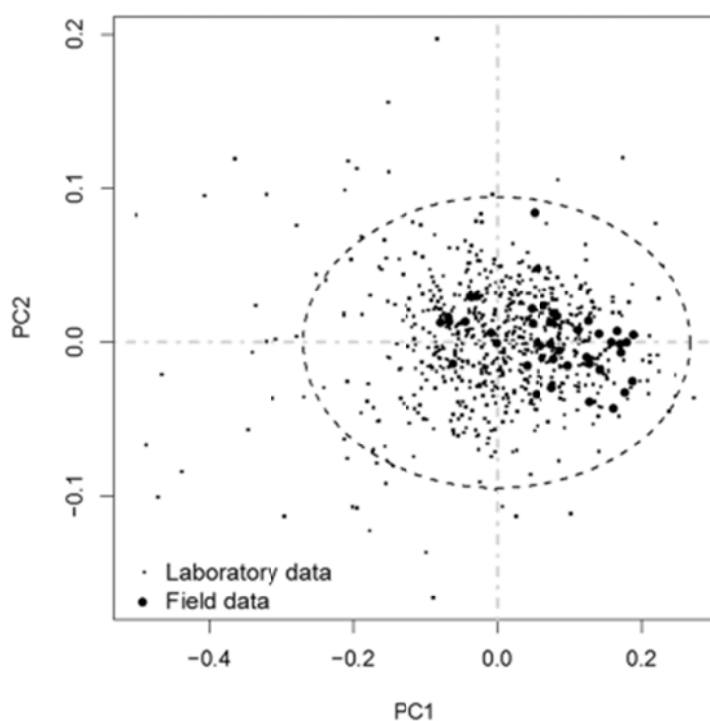


Figure 6. Two-dimensional principal components analysis score plots of laboratory and field samples. Samples on the center of the coordinate systems are laboratory data ($n=700$) used

to train the predictive model (dots), while the others are field data (n=43) where the model was validated (solid circles). Dashed line represents CI 95% of laboratory samples.

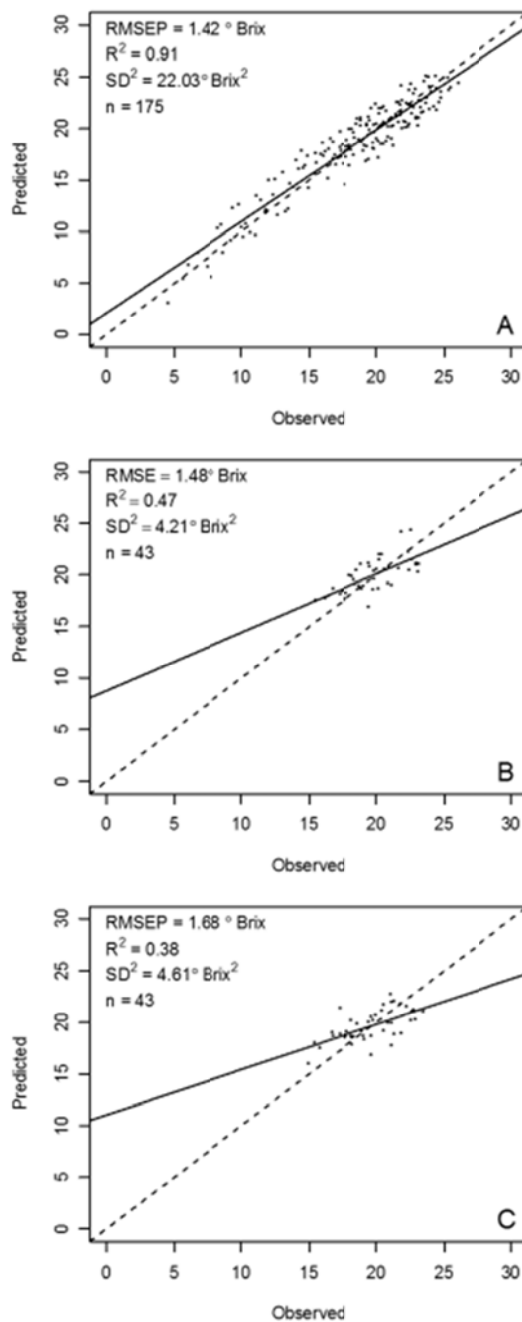


Figure 7. Observed vs. predicted plot obtained with the predicted model in the laboratory prediction set (A) and in the field set (C). In plot B, 43 samples of the laboratory prediction set within the TSS range of the field set were randomly sampled to compare laboratory and field results under similar conditions. Dotted line represents the 1:1 line.

TABLES

Table 1. Statistical parameters of laboratory and in-field sets for total soluble solids ($^{\circ}$ Brix) evaluation.

	Set	N	Min	Mean	Max	SD	
Laboratory	Total	1600	4.60	18.39	29.91	5.11	
	<i>Size of the calibration set</i>	Calibration	50 - 900	5.90 - 4.60	18.63 - 18.65	25.50 - 26.40	4.67 - 4.62
		Prediction	13 - 225	4.60 - 4.80	18.65 - 18.66	27.90 - 27.90	4.63 - 4.68
	<i>Influence of the TSS range</i>	Calibration	700	4.60 - 16.10	18.53 - 20.75	26.90 - 26.40	4.67 - 2.44
		Prediction	175	4.60 - 16.10	18.48 - 20.77	26.90 - 26.90	4.75 - 2.48
	<i>Spectral pre-processing techniques</i>	Calibration	700	4.80	18.63	26.90	4.63
		Prediction	175	4.60	18.53	26.10	4.71
	Epicuticular wax analysis	Total	100	9.82	20.47	26.70	3.15
		Calibration	85	13.22	20.29	26.70	2.94
		Prediction	-	-	-	-	-
In-field	Total	43	15.00	19.63	23.50	2.15	
	Calibration	-	-	-	-	-	
	Prediction	43	15.00	19.63	23.50	2.15	

Number of samples (N), minimum (Min), maximum (Max) and standard deviation (SD). The "total" set includes all recorded samples prior outlier detection while "calibration" and "prediction" sets were the ones used for model training and testing respectively after outlier detection and subsampling.

Table 2. Results of calibration, internal validation and prediction of PLS regression using laboratory dataset with different pre-processing techniques applied to whole-spectrum.

Pre-processing method	Calibration (C)					Cross-Validation (CV)			Prediction (P)				
	R ² C	RMSEC	SEC	Bias	LV	R ² CV	RMSECV	SECV	R ² P	RMSEP	SEP	Bias	RPD
RAW	0.92	1.30	1.30	0.00	15	0.90	1.48	1.48	0.91	1.45	1.52	-0.22	3.10
MC	0.92	1.28	1.28	0.00	16	0.90	1.49	1.49	0.91	1.42	1.40	0.04	3.36
MSC	0.90	1.52	1.52	0.00	11	0.88	1.69	1.69	0.87	1.74	1.73	0.05	2.72
SNV	0.91	1.47	1.47	0.00	13	0.88	1.66	1.66	0.89	1.58	1.53	0.15	3.08
DT1	0.91	1.34	1.34	0.00	12	0.89	1.54	1.54	0.90	1.47	1.53	-0.17	3.08
DT2	0.90	1.40	1.40	0.00	11	0.87	1.59	1.59	0.89	1.53	1.53	0.00	3.08
DT3	0.90	1.45	1.45	0.00	11	0.87	1.66	1.66	0.87	1.65	1.60	0.19	2.94
SG0	0.91	1.36	1.36	0.00	13	0.90	1.47	1.46	0.90	1.46	1.52	-0.17	3.10
SG1	0.91	1.39	1.39	0.00	9	0.89	1.48	1.48	0.88	1.58	1.56	0.07	3.02
SG2	0.91	1.36	1.36	0.00	20	0.88	1.57	1.58	0.87	1.65	1.68	-0.08	2.80

Pre-processing techniques: Mean centering (MC), multiple scatter correction (MSC), standard normal variate (SNV), de-trending using first (DT1), second (DT2) and third order polynomials (DT3) and Savitzky-Golay filter of zero (SG0), first (SG1) and second order (SG2).

Performance measurements: coefficient of determination of calibration (R²C), cross validation (R²CV) and prediction (R²P), root mean square error of calibration (RMSEC), cross validation (RMSECV) and prediction (RMSEP), standard error of calibration (SEC), cross validation (SECV) and prediction (SEP), number of latent variables (LV) and residual predictive deviation (RPD).