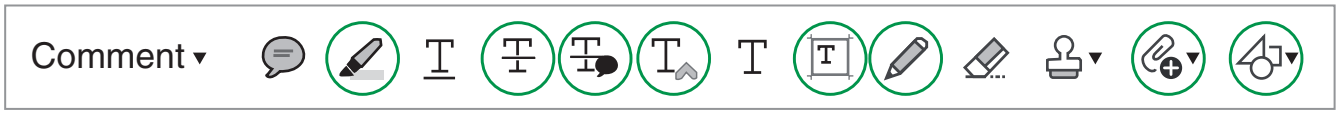
# Page Proof Instructions and Queries

**Journal Title:** Journal of Near Infrared Spectroscopy (JNS)

**Article Number:** 1018759

Thank you for choosing to publish with us. This is your final opportunity to ensure your article will be accurate at publication. Please review your proof carefully and respond to the queries using the circled tools in the image below, which are available in Adobe Reader DC* by clicking **Tools** from the top menu, then clicking **Comment.**

Please use *only* the tools circled in the image, as edits via other tools/methods can be lost during file conversion. For comments, questions, or formatting requests, please use T. Please do *not* use comment bubbles/sticky notes.




*If you do not see these tools, please ensure you have opened this file with Adobe Reader DC, available for free at get.adobe.com/reader or by going to Help > Check for Updates within other versions of Reader. For more detailed instructions, please see us.sagepub.com/ReaderXProofs.

| No. | Query |
|---|---|
| | Please note that we cannot add/amend ORCID iDs for any article at the proof stage. Following ORCID's guidelines, the publisher can include only ORCID iDs that the authors have specifically validated for each manuscript prior to official acceptance for publication. |
| | Please confirm that all author information, including names, affiliations, sequence, and contact details, is correct. |
| | Please review the entire document for typographical errors, mathematical errors, and any other necessary corrections; check headings, tables, and figures. |
| | Please confirm that the Funding and Conflict of Interest statements are accurate. |
| | Please ensure that you have obtained and enclosed all necessary permissions for the reproduction of artistic works, (e.g. illustrations, photographs, charts, maps, other visual material, etc.) not owned by yourself. Please refer to your publishing agreement for further information. |
| | Please note that this proof represents your final opportunity to review your article prior to publication, so please do send all of your changes now. |
| AQ: 1 | Please insert the department/faculty name, wherever applicable, in the affiliation. |
| AQ: 2 | Some values in Table 2 are given in bold format, but an explanation seems to be missing as a table footnote. Kindly check. |

*Original Article*

JNIRS

Journal of Near Infrared Spectroscopy
0(0) 1–12

DOI: 10.1177/09670335211018759
journals.sagepub.com/home/jns
SAGE

# Near infrared spectroscopy determination of chemical and sensory properties in tomato

**Dong Sun[1], Jordi Cruz[2], Manel Alcalà[1], Roser R del Castillo[3,4], Silvia Sans[3] and Joan Casals[3,4]**

**Abstract**

Fast and massive characterization of quality attributes in tomatoes is a necessary step toward its improvement; for sensory attributes this process is time-consuming and very expensive, which causes its absence in routine phenotpying. We aimed to assess the feasibility of near-infrared spectroscopy as a fast and economical tool to predict both the chemical and sensory properties of tomatoes. We built partial least squares models from spectra recorded from tomato puree and juice in 53 genetically diverse varieties grown in two environments. Samples were divided in calibration (210 samples for chemical traits, 45 samples for sensory traits) and validation sets (60 and 10, respectively) using the Kennard and Stone algorithm. Models from puree spectra gave validation $r^2$ values higher than 0.97 for fructose, glucose, soluble solids content, and dry matter (RSEP% ranged 3.5–5.8). $r^2$ values for sensory properties were lower (ranging 0.702–0.917 for taste-related traits (RSEP%: 9.1–20.0), and 0.009–0.849 for texture related traits (RSEP%: 3.6–72.1). For sensory traits such as explosiveness, juiciness, sweetness, acidity, taste intensity, aroma intensity, and mealiness, NIR spectroscopy is potentially useful for scanning large collections of samples to identify likely candidates to select for tomato quality.

**Keywords**

Tomato, fruit quality, sensory analysis, sugars, chemometrics, phenomics, glucose, fructose, soluble solids, dry matter

Received 24 August 2020; accepted 9 April 2021

## Introduction

Worldwide, the tomato (*Solanum lycopersicum* L.) is the second largest vegetable crop, with 182.3 million tons produced for fresh vegetable markets and food processing industries.[1] In recent years, consumer demand in developed countries for tomatoes with better sensory qualities has impelled breeders to obtain varieties with high organoleptic and nutritional quality, whether by developing new materials or reintroducing landrace varieties that were no longer widely cultivated.[2] The sensory profile of tomatoes encompasses a complex set of interrelated organoleptic attributes, such as sweetness, acidity, taste and aroma intensities, mealiness, and skin perception. Some of these attributes have been partially correlated with chemical parameters such as dry matter (DM), soluble solids content (SSC), total acidity (TA), the relative abundance of reducing sugars (mainly glucose and fructose) and acids (mainly citric and malic), and the interaction of these traits with the complex matrix of volatile compounds that give tomatoes their distinctive taste.[3–5] However, the complexity of the relations between chemical traits and human senses makes it very difficult to accurately predict organoleptic properties from chemical composition, and sensory analysis remains the most accurate method to characterize these quality traits. Both sensory and chemical analyses are time-consuming and expensive, making them unsuitable for screening large numbers of varieties as required in breeding programs. Thus, breeding has largely ignored organoleptic attributes, resulting in tomatoes that diverge from consumers' ideotypes and in complaints about tomato quality.

Near-infrared (NIR) spectroscopy is a well-established technique for determining the components of fruits and vegetables,[6] and it can be used on easily prepared samples.[7–9] NIR spectroscopy has been

[1]Departament de Química, Facultat de Ciències, Universitat Autònoma de Barcelona, Bellaterra, Spain
[2]Escola Universitària Salesiana de Sarrià (EUSS), Barcelona, Spain [AQ1]
[3]Fundació Miquel Agustí, Campus del Baix Llobregat, Castelldefels, Spain
[4]Department d'Enginyeria Agroalimentària i Biotecnologia, Universitat Politècnica de Catalunya (UPC), Castelldefels, Spain

**Corresponding author:**
Joan Casals, Fundació Miquel Agustí, Campus del Baix Llobregat, Carrer Esteve Terradas 8, Edifici D4, 08860 Castelldefels, Spain.
Email: j.casalsmissio@gmail.com

applied directly on intact tomatoes to discriminate between varieties,[10,11] determine ripening stages,[12,13] quantify color,[14–16] and measure physical properties[17–19] and chemical parameters such as DM, SSC, total sugars, glucose, fructose, ethylene, lycopene, pH, TA, citric acid, and malic acid.[14,15,17,18,20–24] NIR spectroscopy minimizes sample preparations, enabling rapid quality control by aiding decision making in the field and/or during fruit processing.

However, in a breeding program, breeders search for the genotypical value of each material, and the phenotypical variation between individuals due to environmental effects becomes an important source of bias if we consider a single fruit.[25] Thus, given the high intravarietal variability for some organoleptic traits in tomato,[26,27] breeders usually use a homogenized mixture from several fruits (i.e. puree, squeezed fruits using a blender) to represent accessions and approximate their genotypic value. Centrifuging and/or filtering the puree produces a clear liquid (i.e. juice). Various spectroscopy techniques, including NIR, mid-infrared (MIR), and attenuated total reflectance-Fourier transform infrared (ATR-FTIR), have been used to predict chemical parameters in fresh tomato puree,[28] in juice,[29–32] and in commercial tomato products such as concentrate or ketchup.[33,34] Spectra from puree or juice from harvested tomatoes predict chemical parameters better than spectra from intact fruits: whereas $R^2$ values from intact fruits usually fall between 0.5 and 0.9.[14,15,17,18,20,35] $R^2$ values from puree or juice are usually higher than 0.90.[28,31,36] Although obtaining the puree or juice implies a pre-treatment of the sample, increasing the costs of the analysis, it enables to better approach the mean value of the genotype. Nevertheless, the complete optimization of the process is obtained when the NIR models are constructed based on spectra from intact fruits.

Unlike chemical parameters, sensory properties of foods derived from plants have seldom been correlated with NIR spectra.[37–39] In tomatoes few studies have addressed this issue; Peirs et al.[40] correlated aroma, sweetness, acceptance, mealiness, juiciness, and firmness of eight varieties with NIR spectra, and Clément et al.[41] concluded that NIR spectroscopy was useful in determining two factors related to tomato acceptance, maturity (related to color, firmness, TA, pH, DM, and lycopene) and "gustatory index" (related to electrical conductivity, SSC, TA, and pH), although in this work no sensory attributes where directly modelled. Accordingly, the potential of NIR technology to predict the tomato sensory profile has not been fully explored.

In the current study, we aimed to assess the feasibility of NIR spectroscopy for predicting chemical parameters (DM, SSC, glucose, and fructose) and sensory attributes (sweetness, acidity, taste intensity, odor intensity, skin perception, mealiness, firmness, juiciness, and explosiveness) in tomatoes and to explore its use for scanning large collections of samples to identify potential candidates for selection in breeding programs. To broaden the applicability of the models obtained, the experimental design included a set of 53 tomato varieties representing great genetic diversity.

## Materials and methods

### Samples and analysis

To ensure the inclusion of wide genetic diversity for different fruit quality traits (morphology, sugar, and acid content),[42] we obtained samples from 53 different varieties (pure lines and commercial hybrids representative of the different groups described for tomato, including 30 fresh-market varieties, 14 cherry varieties, 7 ripening mutants, and 2 processing varieties). All varieties were grown in two environments: in soil in an open field and in a soilless culture in a plastic greenhouse. Fruits were collected at the red-ripe stage. Chemical analyses were done on all varieties. Due to the sensory panel's limited capacity, descriptive sensory evaluations were done on only 20 phenotypes (i.e. 10 genotypes (five fresh-market/five cherry type) $\times$ two environments); all varieties were tasted in three different sessions. The experimental fields, chemical analyses, and sensory analyses are described in detail elsewhere.[42]

Briefly, regarding sensory traits, trained panelists evaluated four flavor-related traits (sweetness, acidity, taste intensity, odor intensity) in purees and three texture-related traits (skin perception, mealiness, and firmness) in longitudinal slices for the fresh-market type and in fruit halves for the cherry type. For the cherry varieties, two additional texture-related traits important in consumer acceptance (juiciness and explosiveness)[42] were evaluated in whole fruits. The definition of the sensory traits provided by Hongsoongnern and Chambers[43] were used in the training of the panel. With regard to the explosiveness trait, here we refer to the force in which the liquid inside the fruit is released when it is clenched with the teeth. Each attribute was evaluated in a 0 (low intensity) to 10 (high intensity) scale by nine trained panelists. The average was used as the quantitative score for each variable. The reproducibility of the sensory variables was evaluated by means of the standard error of the reference method (SE), as described in section "Data analysis".

For the 10 varieties undergoing sensory analysis, two biological replicates (different fruits) were taken for each combination of variety, environment, and tasting session (sensory set: 120 samples $=$10 varieties $\times$ 2 environments $\times$ 3 session taste $\times$ 2 biological replicates). For the other 43 varieties, 2 biological replicates were taken for each combination of variety and environment (chemical set: 172 samples $=$ 43 varieties $\times$ 2 environments $\times$ 2 biological replicates).

Each replicate (sensory set, and chemical set) consisted in a minimum of four to five fruits. Chemical analyses and NIR spectroscopy were to be done on all 292 (120 + 172) samples.

Regarding the chemical analysis procedure, for each sample, 500 g of the corresponding tomatoes (minimum of four to five fruits per replicate) were pureed in a blender to obtain a homogeneous mixture (i.e. puree). SSC (hand held refractometer, in °Brix) and DM (65°C, 72 h, in %) were determined in the purees. In order to extract the sugars quantitatively, about 30 g of puree were three consecutive times mixed with about 20 mL of water, filtered and brought together to a volume of 100 mL (i.e. juice). Glucose and fructose (in g/100 g fresh weight) were determined in the juices with a high performance liquid chromatography system equipped with a pump (Beckman 110B, Fullerton, CA, USA), an injector (Hewlett Packard Serie 1100, Palo Alto, CA, USA), a refractive index detector (Beckman 156, Fullerton, CA, USA), and a 250 mm × 4.6 mm Luna NH2 column (Phenomenex, Torrance, CA, USA). Each analysis was repeated twice (two technical replicates).

## Spectra measurement

A Foss NIR spectrometer system model 5000 (Foss, Hilleroed, Denmark) was used to record spectra in samples of puree and juice. Given the different physical states of the two types of samples, different spectrometer units were selected for each type. Tomato juice spectra were acquired in transflectance method with a Rapid Content Analyzer (RCA) accessory and also using a gold transflector, which doubles the optical pathway. First, 10 mL of juice was added to a reflectance vessel; then a 1 mm immersion diffuser was placed in the reflectance vessel to fix the optical path length at 2 mm and the vessel was introduced into the analyzer. Tomato puree spectra were acquired in reflectance method with an OptiProbe Analyzer (Foss, Hilleroed, Denmark). This accessory is used for registering highly scattering liquids and slurries, like the tomato puree which is liquid but with a big amount of solids in suspension (flesh, seeds and peels, which cause a lot of differences in the baseline of raw NIR spectra).

First, 10 mL of puree was placed in a plastic tube; then the probe was placed in the tube with a reflecting mirror in front of the probe window to fix the optical path length at 2 mm.

For both types of samples (juice and puree), the spectral range was 1100–2498 nm and the resolution was 2 nm. Scan number was 32 and 3 spectra were acquired for any sample, using the mean spectrum for computations. Background spectra were acquired every 1 h.

## Data analysis

All the chemical and sensory parameters were predicted from puree spectra. Additionally, glucose and fructose content was predicted from juice spectra as well. The juice is a clear liquid but puree contains flesh, seeds and peels, which cause a lot of differences in the baseline of raw NIR spectra. Prediction models, using Partial Least Squares (PLS),[44] were constructed separately for both types of samples.

To reduce peak overlap, noise, and baseline drift, we tried various preprocessing methods: standard normal variate (SNV),[45] Multiplicative Scattering Correction (MSC),[46] Savitzky-Golay smoothing,[47] and first and second-order derivatives with second order polynomial approximation and 7–41 point window size, and combinations of these. In all cases, the pretreated spectra and the property values were mean centered before being submitted to the regression algorithm. To select calibration and validation-prediction sample sets, we used the Kennard and Stone algorithm, adjusting them to ensure similar standard deviations. The sample sample sets were carefully selected and the intercorrelation between Y-variables was studied to avoid undesired indirect prediction. The best models were established from the calibration set, considering common statistics such as correlation ($R^2$) and root mean square error of calibration (RMSEC). To determine the optimal number of factors, we sought the number that explained the largest proportion of Y-variance without significantly decreasing the RMSEC.

Specific spectral bands were selected to cover the absorbance peaks of the target substance and minimize noise and interferences. To identify the optimum spectral range, we applied the jack-knifing criterion,[48] which provides an indication of the wavelengths that most strongly influence the performance of a model. Further, to identify outliers in the calibration set, samples with high values for sample leverage, X-residual variance, or Y-residual variance were considered outliers if the residual values did not decrease after adding one PLS factor to the plot of variance residuals versus sample leverage. After selection of spectral bands and outliers removing in calibration set, the best models were chosen.

For each PLS model, the number of latent variables (LVs) has been determined from a plot of the Y explained variance, RMSEC and the root mean square error of cross validation (RMSECV) against the number of factors, the optimal number of LVs is the one that with a maximum explained variance shows a minimum value of the error and using the minimum number of LVs.

To determine the predictive ability of the PLS models in the validation samples, which were not used to build the models, we considered the coefficient of determination between $X$ and $Y$ ($R^2$),

RMSECV for models with no validation dataset's evaluation, the root mean square error of prediction (RMSEP), and standard error of prediction (RSEP %) of the mean value, as well as the ratio of the standard deviation of the original data to the standard error of prediction (RPD) and the ratio of the range of the original data to the standard error of prediction (RER).

RMSEC/CV/P, RSEP%, RPD and RER were calculated as follows

$$RMSEC/CV/P = \sqrt{\frac{\sum_{i=1}^{n} \left(y_i^{pred} - y_i^{ref}\right)^2}{n-1}}$$

$$RSEP\% = \frac{RMSEP}{\bar{y}^{ref}} \times 100$$

$$\mathrm{RPD} = \frac{\mathrm{SD}_{y_{ref}}}{\mathrm{RMSEP}}$$

$$RER = \frac{Max\left(y_{i=1,n}^{ref}\right) - Min\left(y_{i=1,n}^{ref}\right)}{RMSEP}$$

where $n$ is the number of samples, $y^{ref}$ and $y^{pred}$ are the reference and predicted value of any sample, and $\mathrm{SD}_{y_{ref}}$ is the standard deviation of the validation reference data.

To develop and calculate PLS models, we used commercially available software (Unscrambler X, Camo ASA; Trondheim, Norway).

The reproducibility of the reference analysis methods was evaluated through the standard error of the reference method (SE), which indicates the uncertainty of the analysis due to the chemical method ($n = 2$ technical replicates) or the panelists ($n = 9$ panelists). SE was calculated as follows

$$SE = \sqrt{\frac{\sum_{i=1}^{n} \left(x_{i,1} - x_{i,2}\right)^2}{2n}}$$

## Results and discussion

### Characteristics of NIR spectra

Raw spectra of both juices and purees showed two absorbance peaks (1386–1578 nm and 1860–2212 nm) and a noise signal from 2350 to 2498 nm (Figure 1(a) and (b)). These two peaks are mostly due to the O–H bond in water molecules, which can decrease the weight of the signal from the target substance. Further, throughout the entire range of spectra, the variation between each spectrum was larger in purees (Figure 1(b)) than in juices (Figure 1(a)), due to interference from solid matters in purees. Pretreatment decreased baseline drift in juices (Figure 1(c)) and puree (Figure 1(d)), although the absorbance of water remained very high and the noise signal from 2350 to 2498 nm also remained.

### Chemical parameters

*Reference data.* Only 270 out of the potential 292 samples (both for puree and juice) could simultaneously be analyzed and their spectrum recorded, because for some genotype*environment combinations there was not enough fruits. For each trait analyzed, 60 samples were selected to serve as a validation set, and the rest were used for calibration. Table 1 reports the chemical composition of the samples in the calibration and validation sets. The environmental and genetic diversity of the accessions studied resulted in large variations in chemical characteristics, covering the entire reported range,[2] thus favoring the likelihood of suitable calibrations and broadening the future applicability of the models.

*PLS models.* Table 2 summarizes the calibration models built with the entire spectrum and with the most useful spectral bands. First, models were constructed based on calculations including the entire range of pretreated spectra from 1100 to 2498 nm. Using this approach, $R^2$ values for the different traits ranged from 0.840 to 0.926 (Table 2).

Next, we adjusted the models by selecting the most useful spectral bands. To determine fructose and glucose in juice samples, we removed the peaks corresponding to water and the high-noise tail, selecting only the spectral band from 2212 to 2310 nm. In puree samples, we used a similar region (2200–2340 nm) to determine DM content; however, to determine SSC, fructose, and glucose, we excluded only the high-noise tail and the second main peak of water (1830–2100 nm) because the puree had a much lower concentration of water than the juices, making it possible to also consider the band around the first peak (1360–1500 nm). Instead of decreasing the quality of the calibration, including this region in the models provided useful information, such as the overtones of O-H in the molecular structure of fructose and glucose.

After the spectral band selection procedure, the quality of calibrations significantly improved (Table 2). $R^2$ values of all models reached values higher than 0.940 and RMSECV values decreased. The improvement was especially strong in analyses of juice samples, where the proportion of water is very high. Additionally, differences between the results obtained with the different pretreatments became smaller.

To further improve the quality of the calibration models, outliers were excluded. No outliers were identified in SSC and DM models. Excluding outliers of fructose and glucose improved the calibration values

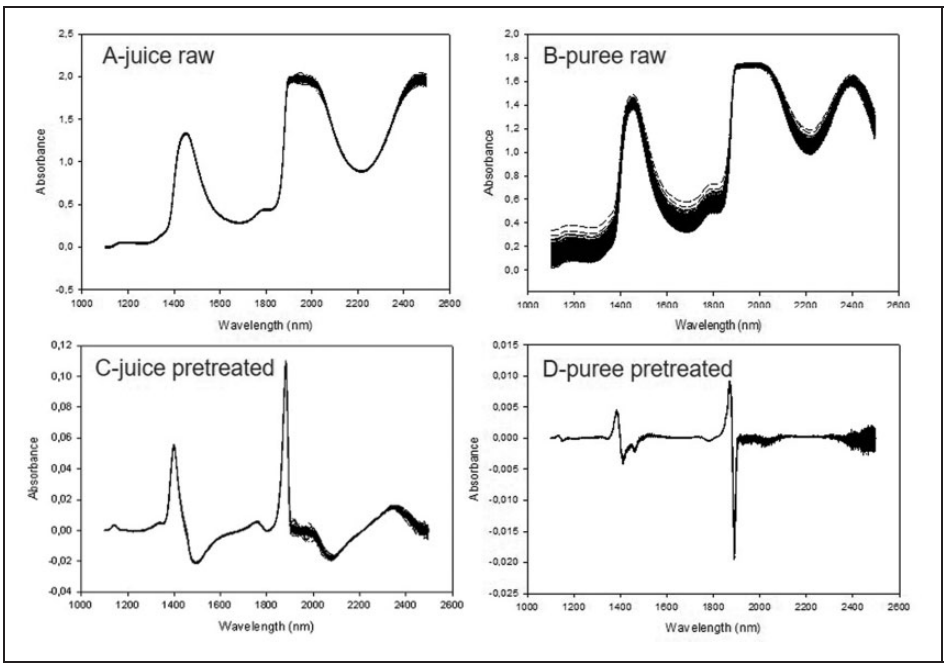



**Figure 1.** Spectra of all tomato samples: (a) raw spectra from juices (270 samples); (b) raw spectra from purees (270 samples); (c) preprocessed spectra from juices (Savitzky–Golay smoothing, first derivative, order 2, 5 points); (d) preprocessed spectra of purees (Savitzky–Golay smoothing, second derivative, order 2, 11 points).

**Table 1.** Chemical composition of the samples in the calibration and validation sets.

| | | Calibration | | Validation | | | |
|---|---|---|---|---|---|---|---|
| Parameter | Range | Mean | SD | Range | Mean | SD | SE |
| Fructose (J) | 4.20–16.23 | 7.28 | 2.51 | 4.78–14.98 | 7.59 | 2.53 | nd |
| Glucose (J) | 3.39–15.48 | 6.93 | 2.58 | 4.40–14.26 | 7.59 | 2.64 | nd |
| Fru ctose (P) | 1.06–3.82 | 1.92 | 0.60 | 1.38–3.52 | 2.03 | 0.62 | 0.041 |
| Glucose (P) | 0.85–3.95 | 1.84 | 0.64 | 1.19–3.56 | 1.99 | 0.65 | 0.042 |
| SSC (P) | 4.20–11.60 | 5.87 | 1.47 | 4.40–10.00 | 6.07 | 1.40 | 0.047 |
| DM (P) | 5.17–11.55 | 7.04 | 1.57 | 5.50–11.00 | 7.43 | 1.61 | 0.048 |

SD: standard deviation, SE: standard error of the reference method, nd: not determined. Units: juice (J), g/L; puree (P), g/100 g.

of $R^2$ and RMSEC slightly for all the models except for glucose in puree (Tables 2 and 3). Figure 2 plots the predicted values versus the reference values from the validation set. Predictions for fructose and glucose from juices led to similar results: $r^2 = 0.970$, RSEP% = 5.8, and $r^2 = 0.971$, and RSEP% = 5.9, respectively. Interestingly, predictions for these parameters from purees were slightly better than those from juices, despite the presence of flesh, seeds and skins; models developed from purees yielded $r^2 = 0.983$ and RSEP% = 4.0 for fructose and $r^2 = 0.982$ and RSEP% = 4.5 for glucose. Better predictions from purees can be explained by two factors: the mean proportion of these sugars in the purees (2 g/100 g) is nearly three times higher than in the juices (0.76 g/100 mL), and the noise from water is lower in the puree spectra than in the juice spectra. Moreover, preparing juices requires several additional steps, so it is more convenient to predict sugars from spectra registered in puree samples. SSC consists mainly of sugars,[2] and predictions of this parameter ($r^2 = 0.972$, RSEP% = 3.9) were nearly as good as those of fructose and glucose. Predictions of DM were also good ($r^2 = 0.976$, RSEP% = 3.5). The values of RPD (5.7–7.7) and RER (21.1–26.4) indicate that the models can make good predictions for all the parameters studied.[7,8]

Table 4 reports statistics summarizing the predictive ability of the PLS models for chemical parameters in juice or puree obtained in the current study and in other reported studies. Although, most of the studies dealing with intact tomatoes or processed products use NIR spectroscopy, the most relevant works with tomato purees or juices have been developed by using MIR technology. For predictions based on spectra recorded in puree, our results for DM ($r^2 = 0.98$) and SSC ($r^2 = 0.97$) are similar to those obtained by Scibisz et al.[28] (DM: $r^2 = 0.96$ and

**Table 2.** Comparison of calibration models built with the whole spectrum and with selected spectral bands. [AQ2]

| Parameter | Pretreatment | Factors | Whole spectrum $R^2$ | Whole spectrum RMSECV | Selected band $R^2$ | Selected band RMSECV | Spectral band (nm) |
|---|---|---|---|---|---|---|---|
| | D1.o2.5p | 7 | 0.819 | 1.073 | **0.974** | **0.411** | |
| Fructose (J) | D2.o2.5p | 7 | 0.838 | 1.015 | 0.970 | 0.439 | 2212–2310 |
| | MSC.D1.cs5.ss2 | 7 | **0.840** | **1.009** | 0.957 | 0.526 | |
| | D1.o2.5p | 7 | 0.832 | 1.062 | 0.949 | 0.587 | |
| Glucose (J) | D2.o2.5p | 7 | 0.823 | 1.090 | **0.963** | **0.498** | 2212–2310 |
| | MSC.D1.cs5.ss2 | 7 | **0.900** | **0.818** | 0.961 | 0.512 | |
| | D1.o2.11p | 3 | **0.874** | **0.215** | 0.938 | 0.151 | |
| Fructose (P) | D2.o2.11p | 3 | 0.784 | 0.282 | **0.941** | **0.148** | 1100–1850, |
| | MSC.D2.cs3.ss3 | 3 | 0.695 | 0.335 | 0.916 | 0.176 | 2094–2228 |
| | D1.o2.11p | 4 | **0.926** | **0.176** | 0.961 | 0.127 | |
| Glucose (P) | D2.o2.11p | 4 | 0.857 | 0.245 | **0.965** | **0.120** | 1100–1850, |
| | MSC.D2.cs3.ss3 | 4 | 0.821 | 0.274 | 0.950 | 0.145 | 2094–2228 |
| | D1.o2.11p | 4 | **0.915** | **0.388** | 0.950 | 0.315 | |
| SSC (P) | D2.o2.11p | 4 | 0.813 | 0.607 | **0.957** | **0.292** | 1100–1834, |
| | MSC.D2.cs3.ss3 | 4 | 0.810 | 0.611 | 0.956 | 0.295 | 2120–2320 |
| | D1.o2.11p | 5 | **0.865** | **0.575** | **0.947** | **0.361** | |
| DM (P) | D2.o2.11p | 5 | 0.787 | 0.722 | 0.918 | 0.448 | 2200–2340 |
| | MSC.D2.cs3.ss3 | 5 | 0.676 | 0.891 | 0.938 | 0.368 | |

The best model for every parameter is written in bold type.
J: juice; P: puree; D: Savitzky–Golay derivative; MSC: multiplicative scattering correction; cs: cap size; o: order; p: points; ss: segment size.
RMSECV units: juice (J), g/L; puree (P), g/100 g.

**Table 3.** Final models for chemical parameters in tomato samples.

| | | | | Calibration | | Validation | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Parameter | Pretreatment | Factors | Outliers | $R^2$ | RMSEC | $r^2$ | RMSEP | RSEP% | RPD | RER |
| Fructose (J) | D1.o2.5p | 7 | 7 | 0.982 | 0.330 | 0.970 | 0.442 | 5.8 | 5.7 | 23.1 |
| Glucose (J) | D2.o2.5p | 7 | 4 | 0.970 | 0.445 | 0.971 | 0.450 | 5.9 | 5.9 | 21.9 |
| Fructose (P) | D2.o2.11p | 3 | 4 | 0.955 | 0.127 | 0.983 | 0.081 | 4.0 | 7.7 | 26.6 |
| Glucose (P) | D2.o2.11p | 4 | 7 | 0.964 | 0.122 | 0.982 | 0.090 | 4.5 | 7.2 | 26.4 |
| SSC (P) | D2.o2.11p | 4 | 0 | 0.957 | 0.292 | 0.972 | 0.237 | 3.9 | 5.9 | 23.7 |
| DM (P) | D1.o2.11p | 5 | 0 | 0.947 | 0.361 | 0.976 | 0.261 | 3.5 | 6.2 | 21.1 |

J: juice, P: puree, D: Savitzky–Golay derivative, o: order, p: points. RMSEC and RMSEP units: juice (J), g/L; puree (P), g/100 g.

SSC: $r^2 = 0.98$) using ATR–FTIR on a set of cherry, fresh-market, and processing tomatoes comprising a total number of samples similar to the present study. Furthermore, their PLS models proved robust when applied to external samples grown in the following year (DM: $r^2 = 0.98$, and SSC: $r^2 = 0.96$). Our results for fructose ($r^2 = 0.98$) and glucose ($r^2 = 0.98$) are slightly better than those obtained by these authors ($r^2 = 0.92$ and $r^2 = 0.96$, respectively), and the same arguments can be raised for the other statistics (RSEP% and RPD).

For predicitions based on spectra recorded in juices, our results for fructose ($r^2 = 0.970$) and glucose ($r^2 = 0.970$) are better than those reported in previous studies using ATR–FTIR ($r^2 < 0.9$).[29–31] Very recently, a study that compares several portable infrared sensing technologies achieves excellent correlations for glucose, fructose and SSC ($r^2 = 0.97$–$0.99$) from both puree and juice.[36]

To ensure the broad applicability of the PLS models we developed, we included materials that represents the phenotypic diversity within cultivated tomatoes: mainly fresh-market and cherry varieties, but also a few ripening mutants (*rin* and *alc*) and processing varieties. This approach increases the diversity of the fruit matrix, thus increasing the noise in the spectra, which should make it difficult to obtain good models. However, using ATR–FTIR, Ibañez et al.[34] found that models built with more diverse groups were generally better than those built with only one group of tomato (processing, cherry or fresh-market varieties). Further, when they applied the models to predict the characteristics of an external sample of processing tomatoes, the models built with more diverse groups were more robust than those built with processing tomatoes alone. They attributed these findings to greater diversity of fruits rather than to a larger number of

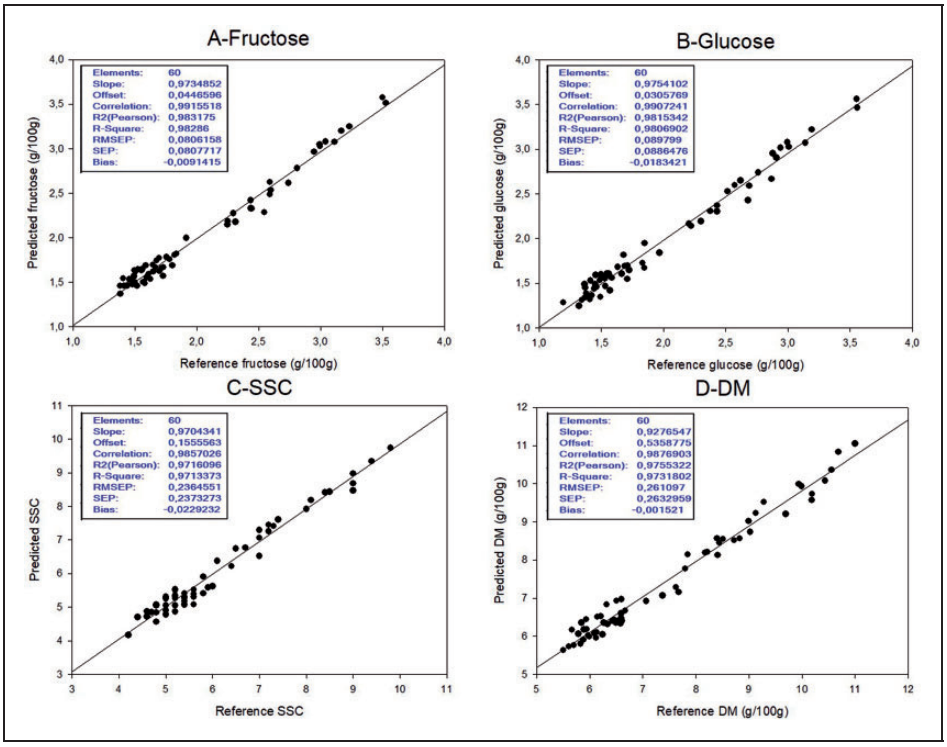



**Figure 2.** Comparisons between reference and NIRS-predicted values for chemical parameters in the validation samples. Spectra recorded in puree: (a) fructose, (b) glucose, (c) soluble solids content (SSC), and (d) dry matter (DM).

samples. These data reinforce the convenience of seeking maximum diversity of fruit types and parameter values in model building.

In summary, the models obtained in the current study from a wide diversity of tomatoes enable good predictions of fructose, glucose, SSC, and DM. The best results were obtained from spectra recorded in puree, which can be easily and rapidly prepared. The RPD values obtained indicate that NIR spectroscopy can be used for quality control of these chemical traits.

## *Sensory attributes*

*Reference data*. Taste tests and NIR spectroscopy could only be done on 26 of the 30 potential samples of cherry tomatoes and 29 of the 30 potential samples of fresh-market tomatoes, because for some genotype*environment combinations there was not enough fruits. The attributes explosiveness and juiciness, distinctive characteristics of cherry tomatoes,[42] were measured only in cherry tomatoes. For these attributes, the calibration set comprised 20 samples and the validation set comprised six samples. For the remaining attributes, measured in both cherry and fresh-market tomatoes (sweetness, taste intensity, odor intensity, mealiness, acidity, firmness and skin perception), the calibration set comprised 45 samples and the validation set comprised 10 samples.

Number of samples employed for sensory attributes is usually lower than these tested for chemical or physical parameters.[35] On the one hand, it is due to the difficulty and time consuming of carry out many taste sessions and on the other hand, to the fact that the individual error associated to each sample increases parallel to the total number of them. So, the number of samples that can be managed in these studies is limited.[49,50]

Table 5 summarizes the sensory panel's ratings (0–10 scale) of the samples in the calibration and validation sets. Scores for most attributes (sweetness, taste intensity, odor intensity, acidity, and firmness) covered a wide range of the scale, though scores for mealiness were all $<6$ and scores for skin perception were all $>4.7$. In cherry tomatoes, scores for explosiveness and juiciness were all $>5$.

*PLS models*. After various pretreatments of raw spectra from purees, we constructed PLS regression models for all sensory parameters considering the whole spectral range (1100–2498 nm) and the spectral range from 2120 to 2320 nm. As discussed before, the high-noise tail from 2320 to 2498 nm and absorption peaks mainly caused by O–H overtones were excluded to avoid excessively large influence of O–H peaks from water in the calibration models. SNV or Savitzky–Golay smoothing with derivative algorithms gave the best calibration results (Table 6). Band selection only improved the results for acidity and explosiveness. Figure 3 plots the values predicted in the validation samples versus the reference values from the calibration set.

**Table 4.** Comparison of dimensionless statistics for predicting chemical parameters in tomatoes with Infrared spectroscopy in studies reported to date.

| Spectrometer | Spectra samples | Type of tomato | Number of samples | | Fructose | Glucose | SSC | DM | Reference |
|---|---|---|---|---|---|---|---|---|---|
| ATR-FTIR 4000–700 $cm^{-1}$ | Juice (centrifuged filtered) | Processing | 245C + 125V | $r^2$ | 0.64 | 0.77 | 0.92 | – | [30] |
| | | | | RSEP% | 7.5 | 7.3 | 4.1 | – | |
| | | | | RPD | – | – | – | – | |
| ATR-FTIR 1880–800 $cm^{-1}$ | Juice (centrifuged filtered) | Fresh-market | 40C + 20V | $r^2$ | 0.88[a] | 0.88[a] | – | – | [29] |
| | | | | RSEP% | – | – | – | – | |
| | | | | RPD | 2.7 [a] | 2.6 [a] | – | – | |
| ATR-FTIR 1500–900 $cm^{-1}$ | Juice (centrifuged supernatant) | Fresh-market Cherry Processing | 245C + 87V | $r^2$ | 0.87 | 0.89 | 0.95 | – | [31] |
| | | | | RSEP% | 14.3 | 12.1 | 4.9 | – | |
| | | | | RPD | 2.8 | 3.6 | 5.5 | – | |
| Transmittance 4000–700 $cm^{-1}$ | Juice (centrifuged supernatant) | Fresh-market Cherry | 400C + 100V Aprox. | $r^2$ | 0.97 | 0.98 | 0.99 | – | [36] |
| | | | | RSEP% | – | – | – | – | |
| | | | | RPD | 4.4 | 5.3 | 8.6 | – | |
| NIR 1000–2500 nm | Juice (+water centrifuged filtered) | Fresh-market Cherry | 205C + 60V | $r^2$ | 0.97 | 0.97 | – | – | This study |
| | | | | RSEP% | 5.8 | 5.9 | – | – | |
| | | | | RPD | 5.7 | 5.9 | – | – | |
| NIR 1000–2500 nm | Puree | Fresh-market Cherry | 205C + 60V | $r^2$ | 0.98 | 0.98 | 0.97 | 0.98 | This study |
| | | | | RSEP% | 4.0 | 4.5 | 3.9 | 3.5 | |
| | | | | RPD | 7.7 | 7.2 | 5.9 | 6.2 | |
| ATR-FTIR 4000–400 $cm^{-1}$ | Puree | Fresh-market Cherry Processing | 260C + 80V | $r^2$ | 0.92 | 0.96 | 0.98 | 0.96 | [28] |
| | | | | RSEP% | 6.8 | 5.1 | 2.9 | 3.8 | |
| | | | | RPD | 2.9 | 4.9 | 6.8 | 5.1 | |
| ATR 4000–700 $cm^{-1}$ | Puree | Fresh-market Cherry | 140C + 40Vaprox | $r^2$ | 0.99 | 0.99 | 0.98 | – | [36] |
| | | | | RSEP% | – | – | – | – | |
| | | | | RPD | 6.2 | 7.7 | 9.6 | – | |

Spectra recorded in puree or filtered puree.
[a]Calculated for calibration samples.

**Table 5.** Sensory properties of the samples in the calibration and validation sets.

| | Calibration | | | Validation | | | |
|---|---|---|---|---|---|---|---|
| Parameter | Range | Mean | SD | Range | Mean | SD | SE |
| Sweetness | 2.48–9.49 | 5.45 | 1.87 | 3.74–9.45 | 6.12 | 1.95 | 0.36 |
| Acidity | 1.79–8.10 | 5.70 | 1.36 | 3.01–8.10 | 5.79 | 1.64 | 0.36 |
| Taste intensity | 2.30–7.70 | 4.86 | 1.18 | 3.53–6.56 | 5.17 | 1.05 | 0.39 |
| Odor intensity | 1.62–7.76 | 3.92 | 1.99 | 1.70–6.51 | 3.86 | 2.16 | 0.30 |
| Mealiness | 1.11–5.96 | 2.73 | 1.38 | 1.19–4.50 | 2.23 | 1.16 | 0.28 |
| Firmness | 1.65–7.96 | 3.54 | 1.67 | 1.89–6.33 | 3.01 | 1.12 | 0.28 |
| Skin perception | 4.71–8.89 | 7.33 | 0.97 | 6.93–8.60 | 7.61 | 0.57 | 0.13 |
| Explosiveness | 5.00–8.53 | 6.57 | 0.90 | 5.57–7.03 | 6.53 | 0.60 | 0.12 |
| Juiciness | 5.83–9.27 | 8.10 | 0.93 | 7.5–9.27 | 8.37 | 0.67 | 0.13 |

Scores in a 0 (low) to 10 (high) intensity scale.
SD: standard deviation; SE: standard error of the reference method.

**Table 6.** Final models of sensory parameters of tomato samples.

| | | | | Calibration | | Validation | | | |
|---|---|---|---|---|---|---|---|---|---|
| Parameter | Taste sample | Factors | Pretreatment | $R^2$ | RMSECV | $r^2$ | RMSEP | RSEP% | RPD |
| Sweetness | Puree | 5 | SNV | 0.712 | 0.990 | 0.917 | 0.559 | 9.1 | 3.5 |
| Acidity | Puree | 4 | D2.o2.7p[a] | 0.326 | 1.532 | 0.702 | 0.970 | 16.8 | 1.7 |
| Taste intensity | Puree | 5 | SNV | 0.342 | 1.115 | 0.762 | 0.672 | 13.0 | 1.6 |
| Odor intensity | Puree | 3 | SNV | 0.642 | 1.278 | 0.875 | 0.773 | 20.0 | 2.8 |
| Mealiness | Slices/halves | 3 | SNV | 0.561 | 0.927 | 0.719 | 0.649 | 29.1 | 1.8 |
| Firmness | Slices/halves | 5 | SNV | 0.335 | 1.819 | 0.381 | 2.170 | 72.1 | 0.5 |
| Skin perception | Slices/halves | 4 | D2.o2.7p | 0.530 | 1.214 | 0.009 | 0.943 | 12.4 | 0.6 |
| Explosiveness | Whole fruit | 6 | D2.o2.7p[a] | 0.653 | 0.326 | 0.849 | 0.232 | 3.6 | 2.6 |
| Juiciness | Whole fruit | 6 | D1.o2.7p | 0.677 | 0.337 | 0.665 | 0.410 | 4.9 | 1.6 |

D: Savitzky–Golay derivative, NV: standard normal variate, o: order, p: points.
[a]Selected band 2120–2320 nm.

For the attributes related to flavor, the best predictions were achieved for sweetness ($r^2 = 0.917$, RPD = 3.5) and odor intensity ($r^2 = 0.875$, RPD = 2.8); the RSEP% was also low for sweetness (9.1%), but not for odor intensity (20%) due to its low mean value in the validation set (3.86). The only texture-related trait for which predictions were acceptable was mealiness ($r^2 = 0.719$, RPD = 1.8). The worst predictions were for firmness ($r^2 = 0.381$, RPD = 0.5) and skin perception ($r^2 = 0.009$, RPD = 0.6); plotting predicted vs. reference values of skin perception resulted in a nearly horizontal straight line, probably because the skin accounts for only a very small part of the entire puree sample. Along these lines, Plans et al.[39] found that NIR spectroscopy was also useless for predicting skin perception in common beans (*Phaseolus vulgaris* L.). For the particular properties of cherry tomatoes, predictions were better for explosiveness ($r^2 = 0.849$, RPD = 2.6) than for juiciness ($r^2 = 0.665$, RPD = 1.6).

To our knowledge, only one other published study has used NIR spectroscopy to predict most of the sensory attributes examined in the present study. Peirs et al.[40] analyzed eight tomato varieties; although no validation results were reported, the best calibration results were obtained for odor ($R^2 = 0.94$) and sweetness ($R^2 = 0.90$), which were also the sensory attributes for which the best correlations were found in our study. Their correlations for mealiness ($R^2 = 0.61$) and juiciness ($R^2 = 0.61$) are not very different from ours. Interestingly, the correlation they found for skin toughness ($R^2 = 0.61$) contrasts with the absence of correlation for skin perception in our study.

NIR spectroscopy predictions are generally worse for sensory attributes than for chemical parameters. The sensory properties of foods depend on both their chemical composition and physical structure. Moreover, interactions between these two aspects also influence panelists' perceptions, which are less precise than measurements with instruments. In an extensive review of NIR application to meat quality, Prieto et al.[51] found that RPD values for sensory attributes were mostly below 1. Values above 2 are scarce.[38,52] Although the RPD values obtained in the current study for sweetness, odor intensity and explosiveness (2.6–3.5) and for acidity, taste intensity, mealiness, and juiciness (1.6–1.8) are not high enough to enable the use of NIR spectroscopy for quality control based on sensory properties, they are

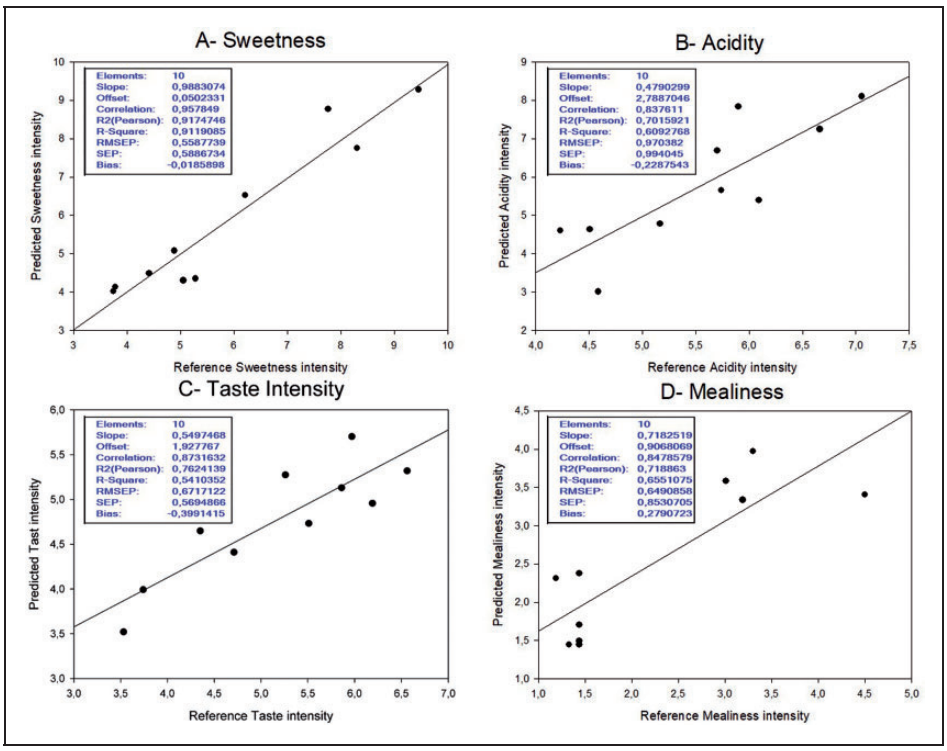



**Figure 3.** Comparisons between reference and NIRS-predicted values for sensory traits in the validation samples: (a) sweetness, (b) acidity, (c) taste intensity, and (d) mealiness.

good enough for screening to discard varieties that are farther from the sensory ideotype. In breeding programs, where large collections must be screened, this approach could drastically reduce the number of samples to be submitted to tasting panels and make it possible for sensory attributes to be included in phenotyping, thus helping breeders meet consumers' demand for better tasting tomato varieties.

## Conclusions

Tomato puree, an easily prepared homogenized mixture of several crushed fruits that mitigates the effects of the high variability between individual fruits, is representative of the qualities of a given accession. Despite the presence of flesh, seeds, and skins, NIR spectra recorded in puree correlated with chemical and organoleptic characteristics and allowed us to predict them. For fructose, glucose, SSC, and DM, $R^2$ values in the validation samples were higher than 0.97 and RPD values were higher than 5.9, indicating that NIR spectroscopy can be used for the quality control of these parameters. Thus, NIR spectroscopy can replace classical chemical determinations in the early stages of breeding programs, allowing large collections to be screened.

As expected, the correlations obtained for sensory properties were lower than those achieved for chemical parameters. However, the $R^2$ and RPD values indicate that NIR spectroscopy can be used to screen for explosiveness and juiciness in cherry varieties and for sweetness, acidity, taste intensity, aroma intensity, and mealiness in a wide variety of tomatoes. Nevertheless, firmness and skin perception cannot be satisfactorily predicted. In summary, NIR spectroscopy can be useful in breeding programs for scanning large collections of samples to identify likely candidates for selection based on sensory characteristics. Further studies should assess the feasibility to use NIR spectroscopy to predict sensory attributes in intact fruits, reducing the costs of analysis.

### Acknowledgements

The authors wish to thank sensory panelists for their work and Semillas Fitó for their technical support, cultivation of the plants, and providing seeds of the materials.

### Authors' contributions

DS: Data curation, Investigation, Validation, Formal analysis; JC: Formal analysis, Methodology; MA: Conceptualization, Supervision; RRC: Data curation, Investigation, Validation; SS: Data curation, Investigation, Validation, Formal analysis; JC: Conceptualization, Funding acquisition, Methodology, Writing – original draft, Writing – review and editing.

### Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Departament d'Agricultura Ramaderia, Pesca i Alimentació de la Generalitat de Catalunya (grant number AAM/259/2013); and the Ministerio de Economía y Competitividad del Gobierno de España (grant number CTQ 2016-79696-P, AEI/FEDER, EU).

## ORCID iD

Joan Casals https://orcid.org/0000-0002-2708-158X

## References

1. FAO. FAOSTAT Food and agricultural commodities statistics, http://faostat.fao.org/ (2018, accessed 18 October 2018).
2. Causse M, Damidaux R and Rousselle P. Traditional and enhanced breeding for quality traits in tomato. In: Razdan MK and Mattoo AK (eds) *Genetic improvement of solanaceous crops. Vol. 2: Tomato*. New Hampshire: Science Publishers, 2007, pp.153–192.
3. Baldwin EA, Goodner K and Plotto A. Interaction of volatiles, sugars, and acids on perception of tomato aroma and flavor descriptors. *J Food Sci* 2008; 73: S294–S307.
4. Tieman D, Bliss P, McIntyre LM, et al. The chemical interactions underlying tomato flavor preferences. *Curr Biol* 2012; 22: 1035–1039.
5. Tieman D, Zhu G, Resende MFR, et al. A chemical genetic roadmap to improved tomato flavor. *Science* 2017; 355: 391–394.
6. Schulz H and Baranska M. Fruits and vegetables. In: Sun D (ed) *Infrared spectroscopy for food quality analysis and control*. San Diego, CA: Elsevier, pp.321–353.
7. Williams PC. Implementation of near-infrared technology. In: Williams PC and Norris KH (ed) *Near-infrared technology in the agricultural and food industries*. St Paul, MN: American Association of Cereal Chemistry, 2001, pp.145–169.
8. Nicolaï BM, Beullens K, Bobelyn E, et al. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review. *Postharvest Biol Technol* 2007; 46: 99–118.
9. Singh CB, Choudhary R, Jayas DS, et al. Wavelet analysis of signals in agriculture and food quality inspection. *Food Bioprocess Technol* 2010; 3: 2–12.
10. Shao Y, Xie C, Jiang L, et al. Discrimination of tomatoes bred by spaceflight mutagenesis using visible/near infrared spectroscopy and chemometrics. *Spectrochim Acta – Part A Mol Biomol Spectrosc* 2015; 140: 431–436.
11. Xie L, Ying Y, Ying T, et al. Discrimination of transgenic tomatoes based on visible/near-infrared spectra. *Anal Chim Acta* 2007; 584: 379–384.
12. Tiwari G, Slaughter DC and Cantwell M. Nondestructive maturity determination in green tomatoes using a handheld visible and near infrared instrument. *Postharvest Biol Technol* 2013; 86: 221–229.
13. Lu H, Wang F, Liu X, et al. Rapid assessment of tomato ripeness using visible near-infrared spectroscopy and machine vision. *Food Anal Methods* 2017; 10: 1721–1726.
14. Ecarnot M, Bączyk P, Tessarotto L, et al. Rapid phenotyping of the tomato fruit model, Micro-Tom, with a portable VIS-NIR spectrometer. *Plant Physiol Biochem* 2013; 70: 159–163.
15. Torres I, Pérez-Marín D, D, la Haba MJ, et al. Fast and accurate quality assessment of Raf tomatoes using NIRS technology. *Postharvest Biol Technol* 2015; 107: 9–15.
16. Acharya UK, Subedi PP and Walsh KB. Robustness of tomato quality evaluation using a portable vis-SWNIRS for dry matter and colour. *Int J Anal Chem* 2017; 2017: 2863454.
17. Sirisomboon P, Tanaka M, Kojima T, et al. Nondestructive estimation of maturity and textural properties on tomato 'momotaro' by near infrared spectroscopy. *J Food Eng* 2012; 112: 218–226.
18. Shao Y, He Y, Gómez AH, et al. Visible/near infrared spectrometric technique for nondestructive assessment of tomato 'heatwave' (*Lycopersicum esculentum*) quality characteristics. *J Food Eng* 2007; 81: 672–678.
19. Camps C and Gilli C. Prediction of local and global tomato texture and quality by FT-NIR spectroscopy and chemometric. *Europ J Hortic Sci* 2017; 82: 126–133.
20. Flores K, Sanchez MT, Perez-Marin D, et al. Feasibility in NIRS instruments for predicting internal quality in intact tomato. *J Food Eng* 2009; 91: 311–318.
21. Xie L and Ying Y. Use of near-infrared spectroscopy and least-squares support vector machine to determine quality change of tomato juice. *J Zhejiang Univ Sci B* 2009; 10: 465–471.
22. Li T, Zhong C, Lou W, et al. Optimization of characteristic wavelengths in prediction of lycopene in tomatoes using near-infrared spectroscopy. *J Food Process Eng* 2017; 40: 1–9.
23. Alenazi MM, Shafiq M, Alsadon AA, et al. Nondestructive assessment of flesh firmness and dietary antioxidants of greenhouse-grown tomato (*Solanum lycopersicum* L.) at different fruit maturity stages. *Saudi J Biol Sci* 2020; 27: 2839–2846.
24. Szuvandzsiev P, Helyes L, Lugasi A, et al. Estimation of antioxidant components of tomato using VIS-NIR reflectance data by handheld portable spectrometer. *Int Agrophysics* 2014; 28: 521–527.
25. Bavay C, Symoneaux R, Maître I, et al. Importance of fruit variability in the assessment of apple quality by sensory evaluation. *Postharvest Biol Technol* 2013; 77: 67–74.
26. Tikunov Y, Lommen A, de Vos CHR, et al. A novel approach for nontargeted data analysis for metabolomics. Large-scale profiling of tomato fruit volatiles. *Plant Physiol* 2005; 139: 1125–1137.
27. Brauss MS, Linforth RST and Taylor AJ. Effect of variety, time of eating, and fruit-to-fruit variation on volatile release during eating of tomato fruits (*Lycopersicon esculentum*). *J Agric Food Chem* 1998; 46: 2287–2292.
28. Ścibisz I, Reich M, Bureau S, et al. Mid-infrared spectroscopy as a tool for rapid determination of internal quality parameters in tomato. *Food Chem* 2011; 125: 1390–1397.

29. Vermeir S, Beullens K, Mészáros P, et al. Sequential injection ATR-FTIR spectroscopy for taste analysis in tomato. *Sens Actuators B Chem* 2009; 137: 715–721.
30. Wilkerson ED, Anthon GE, Barrett DM, et al. Rapid assessment of quality parameters in processing tomatoes using hand-held and benchtop infrared spectrometers and multivariate analysis. *J Agric Food Chem* 2013; 61: 2088–2095.
31. Ibáñez G, Valcárcel M, Cebolla-Cornejo J, et al. FT-MIR determination of taste-related compounds in tomato: a high throughput phenotyping analysis for selection programs. *J Sci Food Agric* 2019; 99: 5140–5148.
32. Jha SN and Matsuoka T. Non-destructive determination of acid-brix ratio of tomato juice using near infrared spectroscopy. *Int J Food Sci Tech* 2004; 39: 425–430.
33. Pedro AMK and Ferreira MMC. Simultaneously calibrating solids, sugars and acidity of tomato products using PLS2 and NIR spectroscopy. *Anal Chim Acta* 2007; 595: 221–227.
34. Saad AG, Pék Z, Szuvandzsiev P, et al. Determination of carotenoids in tomato products using VIS/NIR spectroscopy. *JMBFS* 2017; 7: 27–31.
35. Ibáñez G, Cebolla-Cornejo J, Martí R, et al. Non-destructive determination of taste-related compounds in tomato using NIR spectra. *J Food Eng* 2019; 263: 237–242.
36. Akpolat H, Barineau M, Jackson KA, et al. Portable infrared sensing technology for phenotyping chemical traits in fresh market tomatoes. *LWT* 2020; 124: 109164.
37. Cozzolino D, Smyth HE, Lattey KA, et al. Combining mass spectrometry based electronic nose, visible-near infrared spectroscopy and chemometrics to assess the sensory properties of Australian Riesling wines. *Anal Chim Acta* 2006; 563: 319–324.
38. François IM, Wins H, Buysens S, et al. Predicting sensory attributes of different chicory hybrids using physico-chemical measurements and visible/near infrared spectroscopy. *Postharvest Biol Technol* 2008; 49: 366–373.
39. Plans M, Simó J, Casañas F, et al. Estimating sensory properties of common beans (*Phaseolus vulgaris* L.) by near infrared spectroscopy. *Food Res Int* 2014; 56: 55–62.
40. Peirs A, Desmet M, Nicolaï B, et al. Relations between sensory analysis, instrumental quality and NIR measurements of tomato quality. *Acta Hortic* 2003; 600: 471–477.
41. Clément A, Dorais M and Vernon M. Multivariate approach to the measurement of tomato maturity and gustatory attributes and their rapid assessment by vis-NIR spectroscopy. *J Agric Food Chem* 2008; 56: 1538–1544.
42. Casals J, Rivera A, Sabaté J, et al. Cherry and fresh market tomatoes: differences in chemical, morphological, and sensory traits and their implications for consumer acceptance. *Agronomy* 2018; 9: 9.
43. Hongsoongnern P and Chambers E. A lexicon for texture and flavor characteristics of fresh and processed tomatoes. *J Sens Stud* 2008; 23: 583–599.
44. Romía MB and BernàrdezMA. Multivariate calibration for quantitative analysis. In: Sun D (ed) *Infrared spectroscopy for food quality analysis and control*. San Diego: Elsevier, pp.51–82.
45. Barnes RJ, Dhanoa MS and Lister SJ. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Appl Spectrosc* 1989; 43: 772–777.
46. Isaksson T and Næs T. The effect of multiplicative scatter correction (MSC) and linearity improvement in NIR spectroscopy. *Appl Spectrosc* 1988; 42: 1273–1284.
47. Savitzky A and Golay MJE. Smoothing and differentiation of data by simplified least squares procedures. *Anal Chem* 1964; 36: 1627–1639.
48. Martens H and Martens M. Modified jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR). *Food Qual Prefer* 2000; 11: 5–16.
49. Meilgaard M, Civille G and Carr B. *Sensory evaluation techniques*. 4th ed. New York, NY: CRC Press, 2007.
50. Muñoz AM, Civille GV and Carr BT. *Sensory evaluation in quality control*. Boston: Springer, 1992.
51. Prieto N, Roehe R, Lavín P, et al. Application of near infrared reflectance spectroscopy to predict meat and meat products quality: a review. *Meat Sci* 2009; 83: 175–186.
52. Lapchareonsuk R and Sirisomboon P. Sensory quality evaluation of rice using visible and shortwave near-infrared spectroscopy. *Int J Food Prop* 2015; 18: 1128–1138.