Document downloaded from:

http://hdl.handle.net/10251/158681

This paper must be cited as:

Ibáñez, G.; Cebolla Cornejo, J.; Martí-Renau, R.; Roselló, S.; Valcárcel-Germes, M. (2019). Non-destructive determination of taste-related compounds in tomato using NIR spectra. Journal of Food Engineering. 263:237-242. https://doi.org/10.1016/j.jfoodeng.2019.07.004

The final publication is available at

https://doi.org/10.1016/j.jfoodeng.2019.07.004

Copyright Elsevier

Additional Information

Non-destructive determination of taste-related compounds in tomato using NIR spectra

Ginés IBÁÑEZ[a], Jaime CEBOLLA-CORNEJO[b1], Raúl MARTÍ[b], Salvador ROSELLÓ[a1]*, Mercedes VALCÁRCEL[b]

[a]*Universitat Jaume I, Unidad Mixta de Investigación Mejora de la Calidad Agroalimentaria UJI-UPV. Departament de Ciències Agràries i del Medi Natural, Avda. Sos Baynat s/n, 12071 Castellón, Spain*

[b]*Universitat Politècnica de València. Unidad Mixta de Investigación Mejora de la Calidad Agroalimentaria UJI-UPV. COMAV. Cno. de Vera s/n, 46022 Valencia, Spain*

[1]These authors contributed equally to the work

***Corresponding author:** E-mail: rosello@uji.es, phone: +34 96 487 80 98, fax: +34 96 487 82 16

E-mail addresses: ibanezg@uji.es (G. Ibáñez), jaicecor@btc.upv.es (J. Cebolla-Cornejo), raumarre@upvnet.upv.es (R. Martí), rosello@uji.es (S. Roselló), mervalge@upvnet.upv.es (M. Valcárcel).

ORCID codes: G. Ibáñez: 0000-0002-1787-8587; J. Cebolla-Cornejo: 0000-0002-2607-9920; R. Martí: 0000-0002-3517-1948; S. Roselló: 0000-0002-7733-4178; M. Valcárcel: 0000-0002-9347-1500.

Abstract

Near infrared (NIR) diffuse reflectance was used to predict the contents of taste-related compounds of tomato. Models were obtained for several varietal types including processing tomato, cherry and cocktail tomato, mid-sized tomato and tomato landraces, with a wide range of varieties. Good performance was obtained for the prediction of soluble solids, sugars and acids, considering a non-destructive methodology applied to fruits with different internal structure. Specific models averaged RMSEP (%mean) values lower than 6.1% for SSC, 13.3% for fructose, 14.1% for glucose, 12.7% for citric acid, 13.8% for malic acid and 21.9% for glutamic acid. The performance was dependent on varietal type. General models with a higher number of samples and variation did not improve the performance of specific models. The models obtained, either specific or general, couldn't be extrapolated to external assays and an internal calibration would be required for each assay in order to provide a reliable performance.

Keywords: Fructose, glucose, citric, breeding, *Solanum lycopersicum* L.

1. Introduction

Consumers are often disappointed with the flavor of tomatoes (*Solanum lycopersicum* L.). Several causes explain this situation, ranging from poor genetic material to harvest and handling procedures (Baldwin et al., 2000). Tomato flavor is defined by taste and aroma. Taste is determined by the accumulation of sugars, mainly fructose and glucose, organic acids, mainly citric and malic acid, and the relationship among them. A prominent role for glutamic acid has also been suggested (Bucheli et al., 1999). Increased levels of sugars and acids raised flavor acceptability, though there are maximum levels of acids above which further increases negatively affect consumer acceptability (Malundo et al., 1995). On the other hand, tomato aroma is defined by the accumulation of volatiles. Unlike other crops, aroma in tomato is rather complex, determined by the accumulation and interaction of multiple volatile compounds with none of them holding a prominent role (Baldwin et al., 2000).

Both taste and aroma are also inter-related. Some volatiles associated with fruity or floral notes can enhance the perception of sweetness, and other related to green notes can enhance the perception of sourness (Baldwin et al., 1998). On the other hand, sugars also affect aroma perception. Increased sugar levels enhance the perception of overall, ripe tomato, sweet tomato and tropical aroma notes. Furthermore, increased levels of acids also affect aroma perception. In this case, raising the perception of overall, tropical, ripe tomato and green aroma notes. It also causes a shift from floral and sweet tomato aroma and sweet taste towards bitter and citrus tastes and earthy, green, viney and musty notes (Baldwin et al., 2008).

The development of high-quality tomato productions has become an important objective in order to supply market segments, where some customers value niche products characterized by organoleptic features, giving less importance to the visual quality of the product and willing to pay a premium price (Bazzani and Canavari, 2013). This added value is especially important in the current market context, as after the financial crisis, the level of volatility in tomato prices is especially high and although the prices of tomato for consumers seem to be quite stable, price fluctuations in the chain damage the rest of the agents (Sidhoum and Serra, 2016).

For this purpose, it is necessary to develop new high-quality varieties to offer improved genotypes, to evaluate the growing conditions that optimize the expression of these genotypes and finally to monitor the

production to assure quality standards. Although the recovery of positive alleles involved in the accumulation of volatile compounds that were present in tomato landraces has been proposed (Tieman et al. (2017), the truth is that it is not feasible to analyze the aroma of high quantity of samples at an affordable cost. Accordingly, most emphasis has been placed in the evaluation the accumulation of the sugar and acids, which also play a crucial role in the improvement of tomato flavor. Traditionally, these taste-related compounds have been indirectly measured with gross determinations involving soluble solid contents (SSC) and titratable acidity. But it has been described that sucrose equivalents calculated from the individual accumulation of fructose and glucose is a far better predictor of sweetness and tomato acceptability (Baldwin et al., 1998). And the same applies to organic acids, as it has been reported the positive influence of free acids on sourness (Tandon et al., 2003).

Near-infrared (NIR) spectroscopy offers several advantages over the precise determination of sugars and acids via direct analytical methods based on high pressure liquid chromatography or capillary electrophoresis. It entails an indirect analysis, as NIR data is related to the actual sugar and acid content using chemometrics. Different algorithms have been used for this purpose. In the case of fruit and vegetables the most widely used are least squares regression, LSR, multiple linear regression, MLR, partial least squares, PLS, and principal component regression, PCR (Naes et al., 2002). Among them, PLS is usually preferred over other alternatives for quantification purposes, and PCA as an explorative method (Bureau et al., 2019). In fact, most researches involving spectroscopic data and with NIR and FTIR data choose PLS models (Arendse et al., 2018; Bureau et al., 2019).

The most notable advantage of NIR indirect quantification is that it enables non-destructive indirect determinations, highly valuable in applications that require straightforward, speedy characterization of samples (Blanco and Villarroya, 2002). For this purpose, it has been used in quality analysis of fruits and vegetables. But most works related to taste are targeted to predict gross measurements such as soluble solids contents (SSC) or titratable acidity and using a limited number of varieties (Arendse et al., 2018). Nevertheless, the lacking availability of scientific evidence of the accuracy of these systems is considered a major drawback (Porep et al., 2015).

In this context, several questions have driven the development of the present work. Can efficient NIR PLS regression models be obtained to predict not only SSC but also major sugars and acids in diverse heterogeneous materials with similar characteristics? And in that case, are particular calibrations needed for each assay or general models can be satisfactorily extrapolated?

2. Material and methods

2.1. Plant material

Five sets of samples, each one with a specific material, were used to develop prediction models. The sample sets were configured considering varietal types, usually determined by their size (e.g. cherry and cocktail tomato) and purpose (e.g. processing tomato). The first sample set included 180 samples belonging to eight processing tomato varieties grown with different water and fertilization regimes in Navarra (Spain). The 168 samples from the second sample set were similar but were obtained in Extremadura, a different environment with warmer and sunnier conditions. These samples were obtained during the development of different agronomical studies (Lahoz et al., 2016; Martí et al., 2018). In both cases the fruits had a width in the range of 40-50 mm. The third sample set included 106 samples of 32 varieties of cherry and cocktail tomato (width range 20-35 mm) obtained from local markets. The fourth sample set was more heterogeneous. It represented 108 samples of mid-sized tomatoes (width range 40-82 mm) from 25 varieties including ribbed flat, rounded, plum and cluster tomatoes from commercial and landrace varieties. It was also obtained from local markets. Finally, the fifth sample set included 88 samples of 11 accessions of Spanish tomato landraces (width range 60-120 mm) of the ”Moruno” type, ribbed flat tomatoes similar to the beef type, grown in Albacete (Spain) and kindly provided by Dr. Moreno.

Each specific sample set and a general set with the 650 tomato samples were used for the calculation of models predicting SSC, sugar and acid contents from NIR spectra. In all cases, fully ripe fruits were sampled.

2.2. Acquisition of NIR spectra

All the fruits were washed with water and dried with cellulose tissue. The measurements of the NIR spectrum were carried out at four different and equidistant points in the equatorial peripheral zone of each fruit, as following the four cardinal points, (Hahn, 2002), and measurements were averaged. The spectrum was obtained with a portable NIR spectrometer (Ocean Optics, Dunedin, FL, USA) with an InGaAs detector, covering the range between 902 and 2094 nm, with measurements spaced 6.80 nm, and an optic fibre probe that allowed measurements directly on fruits using diffuse reflectance. The same probe was used for all the varieties independently of the size of the fruit, and if had a space of 20mm between the optical fibre and the edge of the probe. In order to calibrate the equipment a Teflon disk was used as reference, measuring the spectra several times per day.

2.3. Quantification of sugars and acids with capillary electrophoresis

Once the NIR spectra were acquired, the tomatoes were crushed and homogenized. The determination of the soluble solids content was carried out with the obtained tomato juice using a Pocket PAL-α digital refractometer (Atago, Tokyo, Japan). The remaining sample was stored at -80 ° C until the other analytical determinations were made.

The quantification of the reducing sugars fructose and glucose and the organic acids citric, malic and glutamic acids was performed by capillary zone electrophoresis (CZE) with an Agilent 7100 equipment (Agilent Technologies, Waldbronn, Germany) following the method described by Cebolla-Cornejo et al. (2012).

2.4. Chemicals and reagents

Fructose, glucose, citric, malic and glutamic acids, hexadimethrine bromide (HDM), and 2,6-pyridine dicarboxylic acid (PDC), and sodium dodecyl sulfate (SDS) were purchased from Sigma-Aldrich (Steinheim, Germany). Ultrapure water was obtained using a Milli-Q water system (Millipore, Molsheim, France).

2.5. Data analysis

Each sample set was randomly divided into a calibration group (75% of the samples), used to develop the calibration and cross-validation procedures of partial least squares (PLS) regressions, and a validation group (25% of the samples), used to make predictions with the PLS developed. PLS method was selected considering that it is the preferred method for the quantification of sugars and acids in fruits and vegetables using spectroscopic data (Arendse et al., 2018; Bureau et al., 2019).

Before PLS regression, the NIR spectra pre-treatment was performed transforming the diffuse reflectance measured in absorbance (log [1/R]). Subsequently, signal interferences of a multiplicative type, those due to particle size and those associated with changes in wavelength, were eliminated with the SNV correction algorithm (Barnes et al., 1989).

The predictive models were then obtained by PLS regression (Naes et al., 2002). The optimal number of latent variables was calculated using the Venetian blinds cross-validation procedure. Root mean squared errors of calibration (RMSEC) and cross-validation (RMSECV) and the respective coefficients of determination were calculated to check the validity of the results. Minimum RMSECV values and number of latent variables were used as the selection criteria for the number of latent variables to be included in the model. New latent variables were included if they provided a reduction of RMSECV higher than 2%.

At this point, the software provides information regarding outliers in the NIR spectra. The considerations explained by Porep et al. (2015) regarding the identification and removal of outliers were taken into account. Consequently, outliers were removed considering the values of the Hotelling T^2 statistics and the Q residues. In the case of response variables, the values of the normalized residuals (<-3 or >3) and leverage parameters were considered. Then, a definite PLS regression model was recalculated, and the spectra of the samples of the validation group were used to make predictions, calculating the coefficient of determination and root mean squared errors of prediction (RMSEP). RMSEP values were also contextualized using the mean (%mean) of the validation group. Residual prediction deviation (RPD), representing the ratio between the standard deviation of the validation and RMSEP, was calculated to

provide a better comparison between models obtained with different samples. Usually, RPD values higher than 2 represent useful models for classification or quantification (Fearn, 2002).

The reliability of the specific models was studied applying each model to the rest of sample sets. In order to analyse the reliability of general models, five new general models were calculated with four of the sample sets for the calibration and cross-validation and they were later applied to predict the contents using the spectra of the remaining specific sample set.

The pre-treatment of the spectra, PLS regression models, detection of outliers, error parameters and goodness of fit for each model were performed with Matlab v 9.4 (Mathworks Inc, Natick, MA, USA) using the PLS_Toolbox v 8.2.1 module (Eigenvector Research Inc, Wenatchee, WA, USA).

3. Results and discussion

The calibration and validation groups for the specific and general model had similar means and coefficients of variation (Table 1). As expected, the set with cherry and cocktail tomatoes had the highest SSC, and contents of fructose, glucose and citric and malic acids. The group with tomato landraces also had high sugar content, but with much lower citric acid accumulation. In general, a higher level of variation was found for acid contents than for sugars.

Table 1. Statistical parameters of the sample sets used for the calibration and validation of PLS models. N_c: number of samples used for calibration; N_v number of samples used for validation. N: Navarra; E: Extremadura.

		Calibration				Validation			
	Model (N_c/N_v)	Mean	SD	Range	CV (%)	Mean	SD	Range	CV (%)
SSC ºBrix	Processing tomato N (135/45)	4.53	0.56	3.45-6.10	12.3	4.54	0.48	3.60-5.50	10.6
	Processing tomato E (126/42)	4.57	0.47	3.50-5.80	10.3	4.51	0.41	3.65-5.35	9.2
	Cherry&cocktail (80/26)	5.64	1.08	3.95-9.15	19.1	5.51	1.14	3.50-9.00	20.6
	Mid-sized tomato (81/27)	4.30	0.55	2.95-5.65	12.9	4.31	0.55	3.45-5.40	12.7
	Tomato landraces (66/22)	5.42	0.42	4.30-6.20	7.7	5.32	0.53	4.20-6.20	10.0
	General model (487/163)	4.72	0.580	2.95-9.00	16.9	4.79	0.82	3.45-9.15	17.2
Fructose g kg^{-1} fw	Processing tomato N (135/45)	13.42	2.85	5.84-22.42	21.2	12.47	3.12	7.16-19.21	25.0
	Processing tomato E (126/42)	14.46	2.89	8.30-20.04	20.8	14.38	2.44	10.56-18.64	17.0
	Cherry&cocktail (80/26)	19.90	5.63	12.09-38.26	28.3	19.67	5.74	10.35-38.30	29.2
	Mid-sized tomato (81/27)	13.52	3.56	8.15-25.63	26.3	13.86	3.88	8.84-22.75	28.0
	Tomato landraces (66/22)	19.42	2.35	13.85-23.60	12.1	19.02	2.71	13.92-22.95	14.2
	General model (487/163)	15.14	4.36	7.16-38.30	28.8	15.00	4.96	5.09-36.85	33.0
Glucose g kg^{-1} fw	Processing tomato N (135/45)	12.15	2.79	6.10-20.87	22.9	11.67	2.54	7.42-17.16	21.8
	Processing tomato E (126/42)	14.08	2.51	7.60-19.69	18.4	13.88	1.82	10.17-18.67	13.1
	Cherry&cocktail (80/26)	17.30	5.91	9.09-37.25	34.2	17.50	6.61	7.71-38.63	37.8
	Mid-sized tomato (81/27)	11.86	3.34	6.73-22.72	28.1	12.24	3.45	7.11-20.14	28.2
	Tomato landraces (66/22)	17.71	2.67	11.02-22.11	15.1	17.26	3.18	11.13-21.86	18.4
	General model (487/163)	13.89	4.11	6.10-38.63	29.6	13.83	4.40	6.24-35.25	31.9
Citric g kg^{-1} fw	Processing tomato N (135/45)	4.33	0.81	2.13-7.06	18.7	4.18	0.89	2.22-5.69	21.3
	Processing tomato E (126/42)	3.52	0.61	2.03-5.54	17.4	3.50	0.64	2.36-5.22	18.4
	Cherry&cocktail (80/26)	8.61	1.64	5.38-12.38	19.0	8.22	1.53	5.20-10.72	18.7
	Mid-sized tomato (81/27)	5.79	1.94	2.70-14.03	33.6	5.66	1.73	2.98-9.64	30.6
	Tomato landraces (66/22)	4.47	0.81	2.83-6.02	18.2	4.53	0.89	3.13-6.12	19.5
	General model (487/163)	5.05	2.08	2.03-14.03	41.3	5.13	2.20	2.13-11.73	42.9
Malic g kg^{-1} fw	Processing tomato N (135/45)	0.95	0.24	0.37-1.74	25.2	0.91	0.26	0.32-1.27	28.0
	Processing tomato E (126/42)	1.15	0.31	0.48-1.86	26.7	1.21	0.36	0.64-2.00	29.5
	Cherry&cocktail (80/26)	1.42	0.36	0.96-2.55	25.2	1.42	0.54	0.79-3.44	37.6
	Mid-sized tomato (81/27)	1.75	0.65	0.56-4.02	36.9	1.74	0.62	0.56-3.89	35.8
	Tomato landraces (66/22)	1.59	0.50	0.72-2.68	31.7	1.48	0.46	0.82-2.22	30.7
	General model (487/163)	1.26	0.51	0.32-4.02	40.3	1.28	0.51	0.37-3.89	39.7
Glutamic g kg^{-1} fw	Processing tomato N (135/45)	1.75	0.41	0.81-2.77	23.5	1.76	0.45	0.87-2.82	25.7
	Processing tomato E (126/42)	1.03	0.36	0.36-2.35	35.3	1.11	0.46	0.45-2.30	41.6
	Cherry&cocktail (80/26)	1.43	0.91	0.28-4.35	63.7	1.40	0.97	0.32-4.10	69.7
	Mid-sized tomato (81/27)	1.71	0.75	0.70-4.25	44.0	1.64	0.69	0.74-3.70	42.1
	Tomato landraces (66/22)	1.97	0.44	0.92-2.88	22.5	1.87	0.45	0.93-2.65	23.8
	General model (487/163)	1.50	0.66	0.28-4.35	43.9	1.50	0.68	0.43-4.22	45.2

3.1. Prediction models

3.1.1. SSC

Most published works based in non-destructive methods for fruits with thin or thick rind have focused their interest in the indirect quantification of basic parameters such as SSC, titratable acidity and pH (Arendse et al., 2018). In the case of SSC, the performance for prediction varies in each study, with R^2=0.9 and RMSEP=0.4 for apple (Giovanelli et al., 2014), R^2=0.88 and RMSEP=0.46 for pear (Xu et al., 2012), R^2=0.93 and RMSEP=0.62 in peach (Shao et al., 2011) and R^2=0.82 and RMSEP=0.85 in cherry (Escribano et al., 2017).

192 Tomato has also received attention. Even though only one variety has been used in most studies, the
193 performance of NIR based predictions have not been always satisfactory. De Oliveira et al. (2014) tried to
194 develop NIR models predicting SCC in different fruits but concluded that the methodology was not
195 appropriate for fruits with heterogeneous internal structure such as tomato. In fact, their performance for
196 prediction with a single variety was R^2=0.53 and RMSEP=0.53 (%RMSEP=8.9%). Other authors have obtained
197 better performances with their materials. Saad et al. (2015) reached R^2=0.91 and SEP=0.28, again with a
198 single variety. Ecarnot et al. (2013) with the model cultivar "Microtom" obtained a performance for
199 prediction with R^2=0.82 and RMSEP=0.45. Torres et al. (2015) with an obsolete variety, but highly appreciated
200 in the Spanish market, obtaining performances for the prediction with R^2=0.60-0.75 and SEP=0.83-0.65,
201 depending on the hardware used. These last results are similar to the previously obtained by Flores et al.
202 (2009) with the same variety and a validation group of 100 samples (R^2=0.77 and SEP=0.68).

203 In the present work, the performance was highly dependent on the tomato type considered, with R^2_P values
204 for prediction ranging from 0.92 in tomato landraces to 0.51 for processing tomato grown in Navarra (Table
205 2). RMSEP values also ranged from 0.14 to 0.46, which represented 2.7% to 8.4% of the mean value of the
206 validation. RPD values were close to 2, considered a limit to define useful models (Fearn, 2002). These values
207 are similar or even improve those obtained in previous works in tomato or other crops. It is true that the
208 range of variation present in the samples of the calibration model was greater than in other works. This was
209 expected as most works deal with a single variety and in the present work several varieties are present in
210 each specific model. But at the same time this fact also represented a challenge, considering that the
211 interference of the internal structure of tomatoes (pericarp width, number and size of locules, juiciness…)
212 would be much higher as it was much more varied, and differences in internal structure hinder the
213 development of efficient models (de Oliveira et al., 2014).

214 The general model including all the samples had a performance similar to the worse specific model, with
215 R^2_P=0.62 and RMSEP=0.47°Brix, which represents 9.8% of the mean contents (Table 2). Despite being higher,
216 the values obtained with the general model are still similar to those described by other works with a limited

range of varietal variation, and would still be interesting in order to minimize costs in wide screening programs.

Table 2. Performance of NIR based models using partial least squares (PLS) regression predicting contents of taste-related compounds. SSC: soluble solids content; R^2 coefficient of determination; RMSE: root mean squared error; NC: number of samples in the calibration group. NV: number of samples in the validation group; C: calibration; CV: cross-validation; P: prediction; RPD: residual prediction deviation. The number of outliers includes the sum of cases from both the calibration and validation group.

	Model(N_C/N_V)	Outliers	R^2_C	RMSEC	R^2_{CV}	RMSECV	R^2_P	RMSEP	%RMSEP (Mean)	RPD
SSC °Brix	Processing tomato N (135/45)	7	0.89	0.18	0.06	0.69	0.72	0.23	5.1	2.09
	Processing tomato E (126/42)	13	0.81	0.20	0.25	0.43	0.51	0.28	6.2	1.46
	Cherry&cocktail (80/26)	4	0.92	0.31	0.52	0.78	0.87	0.46	8.4	2.48
	Mid-sized tomato (81/27)	10	0.88	0.18	0.64	0.32	0.63	0.34	7.9	1.62
	Tomato landraces (66/22)	4	0.97	0.07	0.33	0.36	0.92	0.14	2.7	3.81
	General model (487/163)	8	0.73	0.41	0.47	0.61	0.62	0.47	9.8	1.74
Fructose g kg^{-1} fw	Processing tomato N (135/45)	8	0.73	1.35	0.08	2.91	0.49	1.95	15.6	1.60
	Processing tomato E (126/42)	8	0.78	1.35	0.35	2.44	0.58	1.69	11.7	1.45
	Cherry&cocktail (80/26)	8	0.86	2.07	0.52	4.04	0.81	2.32	11.8	2.47
	Mid-sized tomato (81/27)	9	0.82	1.47	0.29	3.05	0.32	2.94	21.2	1.32
	Tomato landraces (66/22)	7	0.93	0.64	0.19	2.29	0.82	1.15	6.0	2.36
	General model (487/163)	14	0.58	2.76	0.41	3.31	0.47	3.24	21.6	1.53
Glucose g kg^{-1} fw	Processing tomato N (135/45)	16	0.78	1.12	0.18	2.39	0.42	1.51	13.0	1.68
	Processing tomato E (126/42)	9	0.75	1.21	0.25	2.23	0.50	1.66	12.0	1.09
	Cherry&cocktail (80/26)	6	0.80	2.47	0.58	3.66	0.62	2.87	16.4	2.30
	Mid-sized tomato (81/27)	14	0.84	1.30	0.28	2.94	0.38	2.49	20.4	1.38
	Tomato landraces (66/22)	6	0.91	0.80	0.22	2.54	0.73	1.49	8.7	2.13
	General model (487/163)	14	0.57	2.54	0.41	2.98	0.46	2.92	21.1	1.51
Citric g kg^{-1} fw	Processing tomato N (135/45)	15	0.81	0.28	0.04	0.80	0.71	0.43	10.2	2.08
	Processing tomato E (126/42)	18	0.79	0.25	0.06	0.62	0.65	0.31	8.8	2.08
	Cherry&cocktail (80/26)	5	0.53	1.11	0.22	1.50	0.46	1.17	14.2	1.31
	Mid-sized tomato (81/27)	8	0.68	0.96	0.30	1.47	0.40	1.33	23.5	1.30
	Tomato landraces (66/22)	5	0.94	0.19	0.54	0.57	0.88	0.31	6.9	2.85
	General model (487/163)	23	0.84	0.80	0.73	1.01	0.75	1.00	19.5	2.20
Malic g kg^{-1} fw	Processing tomato N (135/45)	4	0.79	0.11	0.30	0.21	0.71	0.15	16.6	1.73
	Processing tomato E (126/42)	9	0.83	0.12	0.52	0.21	0.73	0.16	13.0	2.28
	Cherry&cocktail (80/26)	7	0.81	0.16	0.48	0.27	0.72	0.18	12.6	3.03
	Mid-sized tomato (81/27)	10	0.80	0.23	0.50	0.37	0.62	0.29	16.6	2.14
	Tomato landraces (66/22)	5	0.96	0.10	0.48	0.37	0.90	0.15	10.3	3.01
	General model (487/163)	23	0.69	0.27	0.53	0.35	0.67	0.28	21.9	1.82
Glutamic g kg^{-1} fw	Processing tomato N (135/45)	12	0.75	0.20	0.17	0.39	0.35	0.25	14.2	1.80
	Processing tomato E (126/42)	14	0.75	0.18	0.35	0.30	0.54	0.24	21.3	1.95
	Cherry&cocktail (80/26)	4	0.85	0.35	0.62	0.57	0.74	0.48	34.3	2.02
	Mid-sized tomato (81/27)	14	0.73	0.32	0.23	0.59	0.26	0.51	31.2	1.35
	Tomato landraces (66/22)	8	0.94	0.10	0.39	0.34	0.81	0.16	8.7	2.77
	General model (487/163)	30	0.51	0.43	0.31	0.53	0.36	0.50	33.1	1.37

The robustness of the models was tested trying to apply each of the specific models obtained to the samples of the rest of sample sets. On the other hand, new general models were calculated with the data of four of the five sample sets and then they were applied to predict the contents of the remaining one. None of the specific models passed the test (Table 3). The highest R^2_P values for the predictions with external assays was

0.20, and RMSEP values ranged from 0.5 to 5.10 in absolute values, and from 10.6% to 100.2% in values contextualized with the mean.

Among the different models applied to predict the rest of assays, the one corresponding to mid-sized tomatoes and the general models had the lowest mean %RMSEP values (20.2% and 20.6% respectively) with absolute values close to 1°Brix (Table 3). In the case of the new general models, R^2_P values were close to 0, with mean %RMSEP values ranging from 11.1%, when it was applied to predict the contents of processing tomato grown in Extremadura to 35.7%, when applied to make predictions with cherry and cocktail tomato, and averaging 20.6% (Table 3).

On the other hand, when the different models were applied to predict the contents of the samples of processing tomato grown either in Navarra or Extremadura and tomato landraces, a lower mean %RMSEP was obtained (17.4, 16.6 and 17.8% respectively). The samples from the cherry and cocktail set and mid-sized tomatoes were more difficult to predict using external calibrations.

Table 3. Performance of NIR based models using partial least squares (PLS) regression for cross-predicting soluble solids content in other assays. SSC: soluble solids content; R^2_P coefficient of determination of the predictions; RMSEP: root mean squared error of the predictions. N: Navarra; E: Extremadura. For each sample set (calibration and validation), the number of samples is indicated.

Model calibration	Model validation	R^2_P	RMSEP °Brix	%RMSEP (Mean)
Processing tomato (N) (168 samples)	Processing tomato (E) (180 samples)	0.031	0.65	14.4
	Cherry&cocktail (106 samples)	0.026	5.10	90.9
	Mid-sized tomato (108 samples)	0.002	4.28	100.2
	Tomato landraces (88 samples)	0.126	0.74	13.7
	General (482 samples)	0.002	3.49	72.7
Processing tomato (E) (180 samples)	Processing tomato (N) (168 samples)	0.001	0.71	15.4
	Cherry&cocktail (106 samples)	0.000	2.65	47.2
	Mid-sized tomato (108 samples)	0.003	2.19	51.3
	Tomato landraces (88 samples)	0.202	1.29	23.9
	General (470 samples)	0.010	1.85	38.5
Cherry&cocktail (106 samples)	Processing tomato (N) (168 samples)	0.008	1.05	22.8
	Processing tomato (E) (180 samples)	0.000	0.95	21.1
	Mid-sized tomato (108 samples)	0.006	2.10	49.2
	Tomato landraces (88 samples)	0.101	0.57	10.6
	General (544 samples)	0.021	1.30	28.3
Mid-sized tomato (108 samples)	Processing tomato (N) (168 samples)	0.013	0.61	13.3
	Processing tomato (E) (180 samples)	0.042	0.50	11.1
	Cherry&cocktail (106 samples)	0.005	1.80	32.1
	Tomato landraces (88 samples)	0.157	1.28	23.7
	General (542 samples)	0.001	1.01	21.0
Tomato landraces (88 samples)	Processing tomato (N) (168 samples)	0.027	1.07	23.3
	Processing tomato (E) (180 samples)	0.121	1.13	25.1
	Cherry&cocktail (106 samples)	0.008	1.88	33.5
	Mid-sized tomato (108 samples)	0.030	1.34	31.4
	General (562 samples)	0.014	1.33	28.3

General models (650-model validation)	Processing tomato (N) (168 samples)	0.082	0.56	12.2
	Processing tomato (E) (180 samples)	0.060	0.50	11.1
	Cherry&cocktail (106 samples)	0.004	2.00	35.7
	Mid-sized tomato (108 samples)	0.000	1.15	26.9
	Tomato landraces (88 samples)	0.046	0.91	16.9

245

246 It is difficult to compare these results with other works, as it is unusual to find the application of the obtained
247 models to external assays. Escribano et al. (2017), in their work with two cherry varieties tried to apply the
248 models of one of the varieties to the other. In that case, the authors concluded that models for SSC did not
249 need to be specific to the variety to be measured to perform adequately. In the present work, neither specific
250 models nor general models were robust enough as to offer reliable predictions in other assays. Even those
251 developed the same varieties and growing conditions but applied to predict contents of samples obtained in
252 a different environment failed to offer a reliable performance. This result emphasizes the need to develop
253 specific calibrations for each assay in order to minimize the error in the indirect predictions.

254

255

256 ***3.1.2. Sugars and acids***

257 The performance of specific PLS models for the prediction of fructose and glucose was highly dependent on
258 the varietal type considered. Mean %RMSEP values of 13.3% and 14.1 were obtained for fructose and glucose
259 respectively, with R^2_P values for prediction ranging from 0.32 to 0.82 (Table 2). Nonetheless, the model for
260 mid-sized tomato offered comparatively high errors, up to 21.2% for fructose and 20.4% for glucose. This
261 group was formed by highly heterogeneous varieties, including flat salad type tomato, plum tomato and
262 cluster tomatoes. The rather heterogeneous internal structure of the varieties would be probably originating
263 a higher level of error in the predictions.

264 The performance of the general model was highly influenced by the worse specific model, with R^2_P for
265 prediction of 0.47 for fructose and 0.46 for glucose and %RMSEP values of 21.6% and 21.1%. As in the case
266 of SSC, general models proved to have low efficacy. Consequently, in this case it would also be recommended
267 to rely on specific models.

Few articles on the determination of specific sugars are available for the quantification of specific compounds, as most published works rely on the determination of basic parameters such as SSC and titratable acidity (Arendse et al., 2018). Among the different limitations of non-destructive NIR spectroscopy for this purpose, the scattering typical of non-transparent media and assignment of NIR bands to specific compounds which absorb in the MIR region have been suggested (Porep et al., 2015). Nonetheless, some data is available. For example, Torres et al. (2015), with a single variety obtained SEP values of 3.8 and 4.0 for fructose and 4.2 for glucose and R^2 values ranging from 0.35 to 0.52, depending on the hardware. Considering mean contents in that work, those values would represent contextualized errors of 20.1%-21.2% for fructose and 19.3% for glucose. Better results were reported by Pedro and Ferreira (2007) with %SEP values of 13.4% for fructose and 11.6% for glucose. In that case, the authors also used diffuse reflectance, but they analyzed samples of tomato concentrate, involving homogenized samples with higher sugar contents. Therefore, a better performance would be logically expected.

The performance of specific models for the indirect quantification of acids was similar to those obtained for sugars, though a worse performance was obtained for glutamic acid. Mean %RMSEP values of 12.7%, 13.8% and 21.9% were obtained for citric, malic and glutamic acid respectively (Table 2). Again, the models for mid-sized tomatoes tended to show a worse performance and the efficiency of the general model was lower than that of the worse specific model. Torres et al. (2015) also modelled citric and malic acid accumulation in their work, obtaining SEP values of 0.81-0.86 and 0.22 respectively, which would represent 18.1-19.2% and 16.5% of the reported mean contents respectively. Most specific models improved these results, while the model for mid-sized tomatoes had similar error levels.

In perspective, mean %RMSEP values obtained in the present work are lower than 15% for fructose, glucose, citric and malic acids, using specific models based on different varieties. These values are considerably good, bearing in mind that they are obtained directly on intact fruits with heterogeneous internal structure. It is true though that the higher level of heterogeneity in fruit internal structure will result in inferior performance, as reported by de Oliveira et al. (2014). That would mean that in order to develop useful models in the industry, the calibration and prediction groups should be formed by fruits with similar structures.

Models based on FT-MIR can be more accurate than those obtained with NIR (Schulz and Baranska, 2009). But the high absorption of MIR radiation in biological tissues entails a low penetration depth, allowing only superficial measurements of a few micrometers (Porep et al. 2015). That means, that MIR indirect measurements require a previous homogenization of the sample and centrifugation of the juices obtained. Undoubtedly, this prior homogenization contributes to a higher accuracy.

The selection of the most appropriate methodology will remain a decision for each industrial/agronomical application. It will be necessary to choose between high-throughput indirect analysis directly on intact fruits with NIR models, with higher error levels and the need to obtained specific calibrations, or obtaining more accurate indirect measurements with general models, but involving a cumbersome pre-processing of samples.

4. Conclusions

One of the main limitations of non-destructive indirect predictions of taste-related compounds based on NIR spectra is that different internal structures of tomatoes can critically affect the performance of the models. In fact, most of the published work available relies on a single tomato variety. Our work proves that it is possible to obtain models with good performance despite this limitation. These models can include several varieties within a specific varietal type and will represent a valuable tool to quantify gross measurements such as soluble solid contents, or even the individual accumulation of fructose, glucose and citric and malic acids. General models can also be obtained, representing a higher number of samples and variability, but their performance would not be better than specific models. More importantly, models must be calibrated for each assay, as the performance of specific or general models to samples obtained in new assays is unacceptable.

Conflict of interest

The authors declare no conflict of interest.

Ackowledgements

This research was performed despite the lack of direct public funding for its development and thanks to the enthusiasm of the authors. The authors thank Dr. Lahoz and Dr. Campillo for providing processing tomato samples and Dr. Moreno for providing samples from tomato landraces. G. Ibañez thanks Universitat Jaume I for funding his pre-doctoral grant (PREDOC/2015/45).

References

Arendse, E., Fawole, O. A., Magwaza, L. S., & Opara, U. L. (2018). Non-destructive prediction of internal and external quality attributes of fruit with thick rind: A review. *Journal of Food Engineering*, *217*, 11-23. https://doi.org/10.1016/j.jfoodeng.2017.08.009

Baldwin, E. A., Scott, J. W., Einstein, M. A., Malundo, T. M. M., Carr, B. T., Shewfelt, R. L., & Tandon, K. S. (1998). Relationship between sensory and instrumental analysis for tomato flavor. Journal of the American Society for Horticultural Science, 123(5), 906-915. https://doi.org/10.21273/JASHS.123.5.906

Baldwin, E. A., Scott, J. W., Shewmaker, C. K., & Schuch, W. (2000). Flavor trivia and tomato aroma: biochemistry and possible mechanisms for control of important aroma components. *HortScience*, *35*(6), 1013-1021. https://doi.org/10.21273/HORTSCI.35.6.1013

Baldwin, E. A., Goodner, K., & Plotto, A. (2008). Interaction of volatiles, sugars, and acids on perception of tomato aroma and flavor descriptors. *Journal of Food Science*, *73*(6), S294-S307. https://doi.org/10.1111/j.1750-3841.2008.00825.x

Barnes, R. J., Dhanoa, M. S., & Lister, S. J. (1989). Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied spectroscopy*, *43*(5), 772-777. https://doi.org/10.1366%2F0003702894202201

Bazzani, C., & Canavari, M. (2013). Forecasting a scenario of the fresh tomato market in Italy and in Germany using the Delphi method. *British Food Journal*, *115*(3), 448-459. https://doi.org/10.1108/00070701311314246

Blanco, M., & Villarroya, I. (2002). NIR spectroscopy: a rapid-response analytical tool. *TrAC Trends in Analytical Chemistry*, *21*(4), 240-250. https://doi.org/10.1016/S0165-9936(02)00404-1

346 Bucheli, P., Voirol, E., de la Torre, R., López, J., Rytz, A., Tanksley, S. D., & Pétiard, V. (1999). Definition of
347 nonvolatile markers for flavor of tomato (Lycopersicon esculentum Mill.) as tools in selection and
348 breeding. *Journal of Agricultural and Food Chemistry*, *47*(2), 659-664. https://doi.org/10.1021/jf980875l

349 Bureau, S., Cozzolino, D., & Clark, C. J. (2019). Contributions of Fourier-transform mid infrared (FT-MIR)
350 spectroscopy to the study of fruit and vegetables: A review. *Postharvest biology and technology*, *148*, 1-
351 14. https://doi.org/10.1016/j.postharvbio.2018.10.003

352 Cebolla-Cornejo, J., Valcárcel, M., Herrero-Martínez, J. M., Roselló, S., & Nuez, F. (2012). High efficiency joint
353 CZE determination of sugars and acids in vegetables and fruits. *Electrophoresis*, *33*(15), 2416-2423.
354 https://doi.org/10.1002/elps.201100640

355 de Oliveira, G. A., Bureau, S., Renard, C. M. G. C., Pereira-Netto, A. B., & de Castilhos, F. (2014). Comparison
356 of NIRS approach for prediction of internal quality traits in three fruit species. *Food Chemistry*, *143*, 223-
357 230. https://doi.org/10.1016/j.foodchem.2013.07.122

358 Ecarnot, M., Bączyk, P., Tessarotto, L., & Chervin, C. (2013). Rapid phenotyping of the tomato fruit model,
359 Micro-Tom, with a portable VIS–NIR spectrometer. *Plant Physiology and Biochemistry*, *70*, 159-163.
360 https://doi.org/10.1016/j.plaphy.2013.05.019

361 Escribano, S., Biasi, W. V., Lerud, R., Slaughter, D. C., & Mitcham, E. J. (2017). Non-destructive prediction of
362 soluble solids and dry matter content using NIR spectroscopy and its relationship with sensory quality in
363 sweet cherries. *Postharvest Biology and Technology*, *128*, 112-120.
364 https://doi.org/10.1016/j.postharvbio.2017.01.016

365 Fearn, T. (2002). Assessing Calibrations: SEP, RPD, RER and R^2. *NIR news*, *13*(6), 12-13.
366 https://doi.org/10.1255%2Fnirn.689

367 Flores, K., Sánchez, M. T., Pérez-Marín, D., Guerrero, J. E., & Garrido-Varo, A. (2009). Feasibility in NIRS
368 instruments for predicting internal quality in intact tomato. *Journal of Food Engineering*, *91*(2), 311-318.
369 https://doi.org/10.1016/j.jfoodeng.2008.09.013

370 Giovanelli, G., Sinelli, N., Beghi, R., Guidetti, R., & Casiraghi, E. (2014). NIR spectroscopy for the optimization
371 of postharvest apple management. *Postharvest Biology and Technology*, *87*, 13-20.
372 https://doi.org/10.1016/j.postharvbio.2013.07.041

373 Hahn, F. (2002). AE—Automation and Emerging Technologies: Multi-spectral prediction of unripe

374 tomatoes. *Biosystems Engineering*, *81*(2), 147-155. https://doi.org/10.1006/bioe.2001.0035

375 Lahoz, I., Leiva-Brondo, M., Martí, R., Macua, J. I., Campillo, C., Roselló, S., & Cebolla-Cornejo, J. (2016).
376 Influence of high lycopene varieties and organic farming on the production and quality of processing
377 tomato. *Scientia Horticulturae*, *204*, 128-137. https://doi.org/10.1016/j.scienta.2016.03.042

378 Malundo, T. M. M., Shewfelt, R. L., & Scott, J. W. (1995). Flavor quality of fresh tomato (Lycopersicon
379 esculentum Mill.) as affected by sugar and acid levels. *Postharvest Biology and Technology*, *6*(1-2), 103-
380 110. https://doi.org/10.1016/0925-5214(94)00052-T

381 Martí, R., Valcárcel, M., Leiva-Brondo, M., Lahoz, I., Campillo, C., Roselló, S., & Cebolla-Cornejo, J. (2018).
382 Influence of controlled deficit irrigation on tomato functional value. *Food Chemistry*, *252*, 250-257.
383 https://doi.org/10.1016/j.foodchem.2018.01.098

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

386 Pedro, A. M., & Ferreira, M. M. (2007). Simultaneously calibrating solids, sugars and acidity of tomato products
387 using PLS2 and NIR spectroscopy. *Analytica Chimica Acta*, *595*(1-2), 221-227.
388 https://doi.org/10.1016/j.aca.2007.03.036

389 Porep, J. U., Kammerer, D. R., & Carle, R. (2015). On-line application of near infrared (NIR) spectroscopy in
390 food production. *Trends in Food Science & Technology*, *46*(2), 211-230.
391 https://doi.org/10.1016/j.tifs.2015.10.002

392 Saad, A., Jha, S. N., Jaiswal, P., Srivastava, N., & Helyes, L. (2016). Non-destructive quality monitoring of
393 stored tomatoes using VIS-NIR spectroscopy. *Engineering in Agriculture, Environment and Food*, *9*(2),
394 158-164. https://doi.org/10.1016/j.eaef.2015.10.004

395 Sidhoum, A., & Serra, T. (2016). Volatility spillovers in the Spanish food marketing chain: The case of
396 tomato. *Agribusiness*, *32*(1), 45-63. https://doi.org/10.1002/agr.21418

397 Shao, Y., Bao, Y., & He, Y. (2011). Visible/near-infrared spectra for linear and nonlinear calibrations: a case
398 to predict soluble solids contents and pH value in peach. *Food and Bioprocess Technology*, *4*(8), 1376-
399 1383. https://doi.org/10.1007/s11947-009-0227-6

400 Schulz, H., & Baranska, M. (2009). Fruits and vegetables. In D. W. Sun (Ed.), *Infrared spectroscopy for food*
401 *quality analysis and control* (pp. 321-353). New York. Elsevier Inc.

Tandon, K. S., Baldwin, E. A., Scott, J. W., & Shewfelt, R. L. (2003). Linking sensory descriptors to volatile and nonvolatile components of fresh tomato flavor. *Journal of Food Science*, *68*(7), 2366-2371. https://doi.org/10.1111/j.1365-2621.2003.tb05774.x

Tieman, D., Zhu, G., Resende, M. F., Lin, T., Nguyen, C., Bies, D., ... & Ikeda, H. (2017). A chemical genetic roadmap to improved tomato flavor. *Science*, *355*(6323), 391-394. http://dx.doi.org/10.1126/science.aal1556

Torres, I., Pérez-Marín, D., De la Haba, M. J., & Sánchez, M. T. (2015). Fast and accurate quality assessment of Raf tomatoes using NIRS technology. *Postharvest Biology and Technology*, *107*, 9-15. https://doi.org/10.1016/j.postharvbio.2015.04.004

Xu, H., Qi, B., Sun, T., Fu, X., & Ying, Y. (2012). Variable selection in visible and near-infrared spectra: Application to on-line determination of sugar content in pears. *Journal of Food Engineering*, *109*(1), 142-147. https://doi.org/10.1016/j.jfoodeng.2011.09.022