

Ripening assessment of 'Ortanique' (*Citrus reticulata* Blanco x *Citrus sinensis* (L) Osbeck) on tree by SW-NIR reflectance spectroscopy-based calibration models

Rosa Pires^{a,*}, Rui Guerra^a, Sandra P. Cruz^a, M. Dulce Antunes^{a,b}, António Brázio^a, Andreia M. Afonso^a, Mariana Daniel^a, Thomas Panagopoulos^c, Isabel Gonçalves^d, Ana M. Cavaco^a

^a CEOT, Universidade do Algarve, Campus de Gambelas, 8005-189 Faro, Portugal

^b MED, Instituto Mediterrâneo Para a Agricultura, Ambiente e Desenvolvimento, Universidade do Algarve, Campus de Gambelas, 8005-189 Faro, Portugal

^c CinTurs, Centro de Investigação em Turismo, Sustentabilidade e Bem-Estar, Universidade do Algarve, Campus de Gambelas, 8005-189 Faro, Portugal

^d CACIAL CRL, Vale da Venda Apartado 128, 8001-902 Faro, Portugal

ARTICLE INFO

Keywords:

'Algarve Citrus'
Ripening
Optimal harvest date
Model recalibration
Multi-orchard
Non-destructive

ABSTRACT

The aim of this study was the non-destructive assessment of 'Ortanique' (*Citrus reticulata* Blanco x *Citrus sinensis* (L) Osbeck) ripening, based on the prediction of internal quality attributes (IQA) by short-wave near-infrared reflectance spectroscopy (SW-NIRS) calibration models. Spectra from fruit of 50 trees located in two different orchards, were acquired on tree using a customized portable visible near-infrared (vis-NIR) system. Partial least squares (PLS) was used to build the various IQA calibration models. The models were tested through internal validation (IV) and external validation (EV). Generally, the IV results were always superior to those of EV: regarding IV, a high regression coefficient (R^2) and low root mean square error of prediction (RMSEP) were achieved, revealing a good predictive performance for juice pH ($R^2 = 0.80$; RMSEP = 0.10; SDR = 2.23), soluble solids content (SSC) ($R^2 = 0.79$; RMSEP = 0.75 %; SDR = 2.27), titratable acidity (TA) ($R^2 = 0.73$; RMSEP = 0.24 % citric acid; SDR = 1.94) and the maturation index (MI) ($R^2 = 0.80$; RMSEP = 1.38; SDR = 2.2). The best EV predictions were obtained for TA ($R^2 = 0.69$; RMSEP = 0.38 % citric acid; SDR = 1.24), and MI ($R^2 = 0.69$; RMSEP = 2.07; SDR = 1.49). Calibration models for glucose, fructose and sucrose showed medium-coarse predictions for both validation strategies. A detailed investigation of MI models was performed, to understand the causes of their poor EV results. In the context of EV, model updating strategies were explored by using some validation samples to improve the calibration model. The methods of bias correction and spiking were tested, showing a clear improvement in the predictions.

1. Introduction

'Ortanique' (*Citrus reticulata* Blanco x *Citrus sinensis* (L) Osbeck) is a valuable tangor mandarin included in the Portuguese Protected Geographical Indication (PGI) 'Algarve Citrus', and highly appreciated by consumers. As a non-climacteric fruit, 'Ortanique' is only harvested at its optimal edible ripening stage (Ladaniya, 2008). To comply with the PGI norm, additionally to the proper external appearance and sanitary conditions of the fruit, the decision of its optimal harvest date (OHD) is based on minima values of the following internal quality attributes (IQA): maturation index ($MI \geq 11$) [ratio of soluble solids

content (SSC) to titratable acidity (TA)], SSC (≥ 11 %) and fruit juice percentage (≥ 35 %) (Uniprofrutal, 2003).

The current approach to establish the OHD of 'Ortanique', similarly to other citrus varieties, is to collect a small number of fruit from each orchard at the beginning of the harvest season and to determine the mandatory IQA by standard destructive and time-consuming methods (Ladaniya, 2008). However, due to the large variability inter- and intra-tree within each orchard, as shown for 'Ortanique' (Pires et al., 2019), or between orchards, as reported for two varieties of sweet orange (*Citrus sinensis* (L) Osbeck) (Cavaco et al., 2021a), the current number of fruit measured seldomly accounts for the whole orchards'

* Corresponding author.

E-mail address: rmpires@ualg.pt (R. Pires).

<https://doi.org/10.1016/j.postharvbio.2021.111750>

Received 8 June 2021; Received in revised form 26 September 2021; Accepted 27 September 2021

Available online 6 October 2021

0925-5214/© 2021 Elsevier B.V. All rights reserved.

heterogeneity. Inadequate fruit sampling can result in a limited assessment of their ripening, heterogenous quality, a deficient OHD management and a weak traceability in the citrus supply chain (Canavari et al., 2010; Haleem et al., 2019).

Non-destructive methods of monitoring fruit, such as visible near-infrared spectroscopy (vis-NIRS), allow a fast and extensive prediction of IQA, through specific calibration models built for various fruit species as reviewed by (Nicolai et al., 2007; Arendse et al., 2018; Pasquini, 2018; Walsh et al., 2020a), including citrus (Magwaza et al., 2012; Cavaco et al., 2021b). However, most IQA calibration models published rely on samples collected and assessed under controlled temperature conditions in the laboratory, as reported for mandarin (Guthrie et al., 2005), several varieties of sweet orange, namely, 'Newhall' (Liu et al., 2015), 'Valencia' (Bizzani et al., 2017; Borba et al., 2021; Ncama et al., 2017), and 'Star Ruby' grapefruit (*Citrus x paradisi* Macfad) (Ncama et al., 2017). Despite the availability of portable spectrometers, only a few studies have focused on their capacity to assess the IQA of oranges and mandarins on tree, perhaps due to the complexities involved and the performance deterioration of calibration models. Those reported, present in general, models with a moderate predictive performance for the various IQA, namely SSC in five cultivars of sweet orange from California (Zude et al., 2008), pericarp thickness, SSC, MI, TA and BrimA in 'Clemenvilla' mandarin and 'Powell Navel' orange (Sánchez et al., 2013a; Torres et al., 2017, 2019), SSC and maximum penetration force in 'Powell Summer' orange (Sánchez et al., 2013b), SSC, MI and juice pH in 'Newhall' orange (Cavaco et al., 2018), and for SSC, chlorophyll and sugar content of several citrus cultivars (Stivastava and Sadistap, 2018).

Additionally, only some of these studies comply with the critical calibration for model robustness to incorporate vis-NIRS in the supply chain daily routines, through a stringent multi-year and multi-orchard validation, such as previously reported for mandarin (Guthrie et al., 2005; Magwaza et al., 2014), and oranges (Zude et al., 2008; Cavaco et al., 2018). In some cases, the recalibration of models through small batches from different orchards or seasons has been proposed, through different techniques, in order to increase the predictive ability of the initial calibration models. This has been done for oranges (Zude et al., 2008; Ncama et al., 2020) as well as other fruit (Mishra et al., 2021). However, the usefulness and impact of model recalibration in real-time conditions through the harvest season need further studies. To our best knowledge, no similar studies were published concerning 'Ortanique'.

The three objectives of this study were: i) to establish the feasibility of SW-NIRS in the non-destructive assessment on tree of several IQAs in 'Ortanique', including the usual attributes, SSC, TA and MI, in addition to sugars and acids; ii) to investigate if that assessment could be performed throughout ripening; iii) to determine the robustness of the calibration models in internal validation and external validation, and to understand how model recalibration (bias correction and spiking) could improve EV results in the follow-up of an orchard along the season.

2. Materials and methods

2.1. Fruit

Tangors (*Citrus reticulata* Blanco x *Citrus sinensis* (L.) Osbeck 'Ortanique') were randomly picked at the eye level (circa 1.60 m height) of the canopy from 50 geo-referenced trees from two commercial orchards, previously selected in order to cover the maximum plot area (25 trees per orchard). Orchards were located in Vale Judeu (Loulé) (37° 7' 8,52" N, 8° 6' 56,19" W; elevation 44 m) and in Gambelas (Faro) (37° 3' 24,22" N, 7° 58' 7,19" W; elevation 16 m), with a tree spacing of 5 × 3 and 6 × 3 × 3, respectively. Although Algarve region has a hot summer Mediterranean climate, and both orchards are located at the same distance from the coast, these two sites present distinct edaphoclimatic conditions, such as, more extreme temperatures (maxima and minima) in Vale Judeu, and different soil types. The commercial season for this variety is February-April (<https://cacial.com/calendario-de-comercializacao-de-citros/>), and fruit were sampled in 2016/2017 harvest season, namely, on November 28th (sampling 1), 2016, January 24th (sampling 2), February 15th (sampling 3), March 14th (sampling 4) and May 9th of 2017 (sampling 5), totaling 125 fruit per orchard (5 fruit per tree). The annual average solar radiation in Algarve (>1918 kW h/m²) presents a variation between 7–12 h/d of effective sunshine in winter and summer, respectively (Atlas do Ambiente, 1995).

citros/), and fruit were sampled in 2016/2017 harvest season, namely, on November 28th (sampling 1), 2016, January 24th (sampling 2), February 15th (sampling 3), March 14th (sampling 4) and May 9th of 2017 (sampling 5), totaling 125 fruit per orchard (5 fruit per tree). The annual average solar radiation in Algarve (>1918 kW h/m²) presents a variation between 7–12 h/d of effective sunshine in winter and summer, respectively (Atlas do Ambiente, 1995).

2.2. Spectroscopy under field conditions

One spectrum per fruit was acquired on-tree, on the same day in both orchards, using a customized portable vis-NIRS system (Jaz, Ocean Optics, USA), working in the full range 659–1136 nm (restricted to narrower ranges in data analysis, as explained in Sec. 2.4) and with an average optical resolution of 3.5 nm, according to Cavaco et al. (2018). The system consisted of a spectrometer, a tungsten light source and a custom-made interactance probe with a bifurcated optic fiber. This interactance probe included a central receptor fiber, connected to the spectrometer and six emitting-fibers connected to the light-source, which are distributed in a 5 mm circle around the receiving fiber. The optics interaction took place only within a 10 mm circle, which is small enough to adapt almost perfectly to the fruit, which never had diameters below 65 mm. A Spectralon disc (WS-1, Ocean Optics, USA) was used as absolute reference. An aluminum cup was adapted to acquire the spectra from fruit on tree without the interference of ambient light and to avoid its detachment from the branch. Each spectrum acquired was an average of 40 scans, and a convenient integration time was chosen to obtain a proper signal/noise ratio. The large number of scans was needed to reduce the noise above 950 nm, where the detector has low quantum efficiency. These scans were performed continuously, with the probe in contact with the same point on the fruit. For the spectra acquisition, a specific software was used (Spectra Suite, Ocean Optics, USA). After spectra acquisition, fruit temperature was determined with an infra-red handheld digital thermometer (Fluke FoodPro Plus thermometer, USA). Then, each fruit was collected, numbered and taken to the laboratory for further analysis. The purpose of temperature measurements was simply to verify that it was within acceptable limits, since reflectance spectra are affected by temperature (Golic et al., 2003; Peirs et al., 2003). Unpublished studies carried out in pear, showed that the prediction models worsen much below 10 °C and above 40 °C. In this study, the fruit temperatures were all between 17 °C and 31 °C, which guarantees that the temperature effect in the models is minimal.

2.3. Quality attributes

After equilibration at 20 ± 2 °C, each fruit was weighed and its maximal equatorial diameter (size) determined with a caliper. Firmness was determined by compression of the fruit with a texture meter (Chatillon Force TCD200, Digital Force Gauge Dfis50, Ametek, USA) fitted with a 65 mm diameter flat plate at 1 mm/s speed to compress fruit for 10 mm from contact point, according to Singh and Reddy (2006). Each fruit was then squeezed individually in an orange automatic squeezer and filtered. The total juice percentage (w/w), SSC, TA and juice pH were determined using the usual standard procedures (Ladaniya, 2008). A fraction of the juice was used to measure SSC with a digital refractometer (Atago Co. Ltd., Tokyo, Japan). Juice pH was determined with a digital pH meter (TitroLine® 6000, SI Analytics GmbH, Germany). TA was determined in 5 mL orange juice diluted with 5 mL distilled water subjected to volumetric alkaline titration with 0.1 N NaOH until a pH of 8.2, using an automatic potentiometric titrator (TitroLine® 6000, SI Analytics GmbH, Germany), and was expressed as mass percentage of citric acid per 100 mL juice (% citric acid). MI was calculated from the ratio SSC/TA. Juice samples of each fruit were immediately frozen in liquid nitrogen and kept at -80 °C until assay. Soluble sugars (glucose, fructose and sucrose) and L-ascorbic acid were assayed enzymatically with the commercial kits K-SUFRG-Sucrose/D-Fructose/D-Glucose and

K-ASCO from Megazyme® (Megazyme International Ireland Limited, Ireland), respectively.

2.4. Data analysis

The effect of sampling time/date and orchard on the various IQA measured was assessed by an ANOVA of Repeated Measures (Type III) and Tukey HSD post hoc test for a significance level of $p < 0.05$ (SPSS 24, IBM SPSS Inc., NY, USA). A two independent Student's *t*-test was performed in the comparative analysis of the means of the aforementioned parameters

for a significance level of $p < 0.05$. Mann-Whitney nonparametric test was performed to compare the PAR measured in the two orchards and a one-way ANOVA was performed to compare the three-calibration datasets used to build the models (SPSS 24, IBM SPSS Inc., NY, USA).

All the other data analyses were performed in Matlab® 2019a (The MathWorks, Inc., Natick, MA, USA) and in PLS Toolbox for MATLAB, version 8.7, 2019 (Eigenvector Research, Manson, WA, USA). All the calculations were automated, through the inclusion of fundamental PLS Toolbox functions into scripts written by the laboratory team.

Table 1

Summary of the results obtained in IV and EV for MI, TA, SSC, firmness, juice pH, sucrose, total sugars and ascorbic acid calibration models. G stands for Gambelas and VJ for Vale Judeu.

IQA	CAL→VAL	nLV	Preprocess	RMSECV	RMSEC	RMSEP	SDR	R ²	Bias	Slope
MI SEL=0.2	VJ→G	3	ABS + SG (25,0,0) + detrending + SG (25,2,1)	1.40	1.16	2.61	1.56	0.71	-0.98	0.49
	VJ→G	5	ABS + SNV	1.32	1.04	2.93	1.39	0.74	-1.79	0.51
	G→VJ	3	ABS + SG (25,0,0) + detrending + SG (25,2,1)	2.01	1.77	2.07	1.49	0.69	1.14	0.73
	G→VJ	4	ABS + SNV	1.89	1.76	3.08	1.00	0.63	2.06	0.97
	VJ→VJ (*)	5	ABS + SNV	1.35	1.38	1.38	2.25	0.80	0.01	0.84
	G→G (*)	4	ABS + SNV	2.00	2.05	2.05	1.92	0.73	0.12	0.78
TA (% citric acid) SEL=0.03 g/ 100 mL	(G+VJ) → (G+VJ)	5	ABS + SNV	1.78	1.74	1.74	2.11	0.78	-0.04	0.78
	VJ→G	2	ABS + CR (700,1100) + SNV + SPLINE (30,5,2)	0.25	0.24	0.38	1.24	0.69	-0.28	0.68
	G→VJ	5	ABS + CR (700,1100) + SNV + SPLINE (30,5,2)	0.27	0.24	0.39	1.17	0.69	0.30	0.80
	VJ→VJ	2	ABS + SNV	0.26	0.21	0.27	1.80	0.69	-0.01	0.71
	G→G	2	ABS + SNV	0.23	0.21	0.24	1.94	0.73	0.00	0.74
	(G+VJ) → (G+VJ)	2	ABS + SNV	0.26	0.23	0.25	1.93	0.73	0.01	0.75
SSC (%) SEL=0.1 %	VJ→G	6	ABS + CR (700,900) + SNV + NORRIS (80,15,2)	0.87	0.58	1.61	0.76	0.43	-1.17	0.77
	G→VJ	3	ABS + CR (700,900) + SNV + NORRIS (80,15,2)	0.88	0.75	1.83	0.88	0.60	1.53	0.57
	VJ→VJ	6	ABS + SNV	0.91	0.58	0.75	2.27	0.79	-0.03	0.84
	G→G	4	ABS + SNV	0.85	0.69	0.83	1.86	0.72	0.00	0.77
	(G+VJ) → (G+VJ)	7	ABS + SNV	0.94	0.63	0.93	1.79	0.69	-0.03	0.73
	VJ→G	5		30.68	25.70	29.44	1.22	0.38	7.83	0.45
Firmness (N) SEL=n.a.	G→VJ	5		25.19	20.22	31.94	1.24	0.34	0.24	0.35
	VJ→VJ	2	CR (700,900) + ABS + PSNV (150) + NORRIS (80,15,2)	29.63	26.77	39.76	1.39	0.46	-1.95	0.36
	G→G	6		29.05	19.80	20.91	1.85	0.69	-4.26	0.72
	(G+VJ) → (G+VJ)	5		27.76	24.15	34.43	1.20	0.29	-4.12	0.31
	VJ→G	5		0.16	0.14	0.19	1.27	0.59	0.09	0.83
	G→VJ	4		0.14	0.12	0.20	1.42	0.60	-0.09	0.54
Juice pH SEL=0.07	VJ→VJ	3	CR (700,900) + ABS + PSNV (150) + SG (100,2,2)	0.16	0.15	0.19	1.59	0.63	-0.07	0.59
	G→G	7		0.12	0.09	0.19	1.29	0.40	0.00	0.55
	(G+VJ) → (G+VJ)	4		0.16	0.15	0.10	2.23	0.80	0.01	0.89
	VJ→G	2		1.22	1.15	1.36	1.32	0.43	-0.21	0.43
	G→VJ	4		1.41	1.26	1.22	1.32	0.45	0.24	0.52
	VJ→VJ	5	CR (700,900) + ABS + PSNV (150) + SG (100,2,2)	1.22	1.06	1.04	1.38	0.51	0.13	0.72
Sucrose (g/100 mL juice)	G→G	3		1.36	1.24	1.47	1.30	0.38	-0.31	0.42
	(G+VJ) → (G+VJ)	3		1.24	1.16	1.49	1.23	0.34	-0.14	0.42
	VJ→G	2		1.46	1.40	2.05	1.11	0.37	-1.00	0.35
	G→VJ	2		1.77	1.71	1.86	1.09	0.50	1.19	0.53
	VJ→VJ	2	ABS + CR (730,900) + MSC + SG (100,2,2)	1.49	1.45	1.22	1.62	0.59	0.12	0.58
	G→G	2		1.74	1.60	2.15	1.19	0.30	-0.62	0.25
Total Sugars (g/100 mL juice)	(G+VJ) → (G+VJ)	2		1.69	1.64	1.74	1.33	0.41	-0.04	0.44
	VJ→G	2		0.38	0.35	0.68	0.68	0.16	-0.53	0.17
	G→VJ	3		0.46	0.39	0.57	0.75	0.26	0.44	0.30
	VJ→VJ	2	RANGE (730,1100) + ABS + PSNV (150) + SG (100,2,2)	0.40	0.36	0.30	1.25	0.33	-0.06	0.36
	G→G	4		0.43	0.36	0.46	0.94	0.01	-0.02	0.06
	(G+VJ) → (G+VJ)	3		0.44	0.41	0.30	1.50	0.56	0.00	0.66
Fructose (g/100 mL juice)	VJ→G	5		0.40	0.33	0.51	0.98	0.23	-0.26	0.30
	G→VJ	4		0.47	0.38	0.67	0.71	0.14	0.50	0.18
	VJ→VJ	4	RANGE (730,1100) + ABS + SNV + NORRIS (80,15,2)	0.41	0.35	0.34	1.24	0.35	0.08	0.43
	G→G	4		0.42	0.36	0.50	1.03	0.07	0.00	0.15
	(G+VJ) → (G+VJ)	6		0.41	0.35	0.47	1.28	0.37	-0.05	0.37
	VJ→G	2		2.25	2.19	3.14	0.99	0.28	-1.70	0.36
Ascorbic Acid (mg/100 mL juice)	G→VJ	2		2.60	2.46	2.74	1.29	0.56	1.33	0.43
	VJ→VJ	2	ABS+CR (700,900) + PSNV (150) + NORRIS (80,15,2)	2.29	2.14	2.31	1.62	0.59	-0.09	0.57
	G→G	1		2.78	2.70	2.29	1.29	0.34	0.16	0.29
	(G+VJ) → (G+VJ)	3		2.60	2.36	2.40	1.25	0.37	-0.06	0.49

nLV = number of latent variables, SEL = standard error of laboratory, ABS = Absorbance, CR = continuum removal, NORRIS = Norris derivative (2nd order), SG = Savitzky-Golay derivative (2nd order), RANGE = sets upper and lower wavelength limits, PSNV = piecewise standard normal variate, MSC = multiplicative scatter correction, SNV = standard normal variate, SPLINE = spline smoothing, RMSECV = root mean square of error prediction in cross validation, RMSEC = root mean square of error in calibration, RMSEP = root mean square of error of prediction, SDR = standard deviation ratio, R² = regression coefficient, CAL: calibration set, VAL: validation set. In bold: values for the best calibration model for external validation.

(*) Some differences are found between the IV and EV calibration results for MI due to the different preprocessing and random resampling analysis employed in IV.

2.4.1. Spectra preprocessing

Spectra preprocessing included the elimination of wavelengths below 700 nm to avoid the spectral range covering the chlorophyll peak (680 nm), and wavelengths above 1100 nm because of excessive noise. The final range was further tuned during model calibration. The ratio between chlorophyll and carotenoid content in fruit rind, responsible for the green-orange color, is not a reliable indicator of citrus quality/ripening, depending on various abiotic factors, such as low temperature, and on the production region (Lado et al., 2014).

Spectra were submitted to transformation, through various techniques, namely, standard normal variate (SNV) (Barnes et al., 1989), piecewise SNV (PSNV) (Rinnan et al., 2014), multiplicative scatter correction (MSC) (Martens et al., 1983), Norris derivative (Hopkins, 2001), spline smoothing (SPLINE) (de Boor, 1980) Savitzky-Golay smoothing and derivation (SG) (Savitzky and Golay, 1964) and continuum removal (CR) (Clark and Roush, 1984). For each IQA, the preprocessing transformation yielding the best models were selected and presented in the results section (Table 1).

2.4.2. Outlier elimination in the raw data

During spectra acquisition, signal contamination by ambient light can occur due to misplacement of the probe or the lid that protects the fruit inside the cup. The defective spectra were identified by an abnormal feature around 760–765 nm, visually identified as a dip and which is induced by the absorption of molecular oxygen (the telluric A band of the solar spectrum). The detection of the outliers was performed automatically, by considering the linear interpolation (straight line) between the 755 and the 775 nm, and a smoothed version of the raw data between the same wavelengths (Savitzky-Golay of 31 points and second order polynomial). A given spectrum was considered as contaminated by sunlight if the smoothed curve defined a pronounced dip at 765 nm relatively to the straight line, and if its depth was beyond a predefined threshold value. Fifteen and seven spectra were removed from the dataset collected in VJ and G orchards, respectively.

2.4.3. Calibration models

Calibration models were constructed for the various IQA, namely, SSC, TA, MI, firmness (F), and also for the biochemical analytes, glucose, fructose, sucrose, total sugars, and ascorbic acid content, correlating all fruit spectra and the respective data resulting from the standard analysis through Partial Least Squares (PLS). The complete dataset was split into calibration and validation. The split percentage depends on whether the validation is internal or external, as explained below. The calibration dataset was standardized, and the model built from cross-validation using venetian blinds, with five splits and a blind size of five samples. This means that the calibration dataset is split in five blocks, by assigning 5 successive samples for each block. Once constructed, each calibration model was validated by a set of independent samples (Nicolaï et al., 2007). Then, two main approaches were followed. The first and most commonly used, consists in taking all available spectra and randomly select a fraction of the samples for calibration (80 %) and another for validation (20 %). Thus, the samples in the two sets are distinct, but homogenous because they were acquired from the same database. This process is called internal validation (IV), precisely because of the homogeneity between the calibration and validation sets, and it only allows a proof of concept (Cavaco et al., 2018, 2021b). The second approach is external validation (EV). In this case, the calibration was performed with the dataset from one orchard and the validation performed with the dataset from the other orchard. This procedure allows the degree of heterogeneity between the two datasets to be evaluated and to establish more effectively the accuracy and predictive capacity of the models.

The performance of the models was characterized by the recommended standard parameters (Walsh et al., 2020b): root mean square (RMS) of error in calibration (RMSEC), RMS of error prediction in cross validation (RMSECV), RMS of error of prediction (RMSEP), standard

deviation ratio (SDR) = $\text{STD}(y)/\text{RMSEP}$ where y represents the reference destructive data, correlation coefficient (R), Bias = $\text{mean}(y') - \text{mean}(y)$, where y' represent the predictions, slope of the linear regression y' vs. y (Nicolaï et al., 2007; Walsh et al., 2020a) complemented by other parameter that have shown to be relevant, as previously described by Cavaco et al. (2018), namely $\text{res}(\%) = \text{“resolution”} = \text{RMSEP}/[\text{max}(y) - \text{min}(y)] * 100$, which is a measure of how the prediction error resolves the range of variation of y . The number of latent variables (nLV) of each model was chosen in order to minimize RMSECV.

In order to help structuring the discussion in Section 3, the model performance according to the SDR values will be classified according to the following scheme: class 1 (no prediction): $\text{SDR} < 1$, meaning predictions worse than simply assuming the validation mean for all the samples; class 2 (poor): $1 < \text{SDR} < 1.5$, representing a very weak predictive power; class 3 (coarse): $1.5 < \text{SDR} < 2$, allowing for coarse predictions (high/low values); class 4 (good): $2 < \text{SDR} < 2.5$, which represents already a reasonable predictive capacity.

2.4.4. PLS models with random resampling

Compared to the other parameters, the MI, SSC and TA models were submitted to a more complete statistical treatment in IV, since these parameters define the OHD. The goal of this more detailed treatment was to infer more robust model performance estimates (Table 1). Thus, the PLS models for MI, SSC and TA were built with random resampling. This meant running the models N times. For our calculations we adopted $N = 50$. In each run the validation was performed by choosing randomly 20 % of the fruit as validation set and performing the model calibration through cross-validation in the other 80 % of the fruit. The samples selection was performed without replacement, contrary to the bootstrap approach, in which the same sample could be chosen only once. The process was repeated 50 times, in order to derive statistically more robust results. After each round, the predicted and measured values for the validation dataset were stored successively in the two-column matrix. In the end, the statistical parameters (RMSEP, R^2 , etc.) were calculated for the entire dataset that was concatenated in that matrix, comprising thousands of predictions. This means that the RMSEP obtained after random resampling is the root mean square of the individual RMSEPs obtained in each run.

2.4.5. Models' recalibration

To reduce the prediction error for the EV models, two methods of model recalibration (updating) were investigated: bias correction and spiking (Guerrero et al., 2010). In any case, N samples from the validation set are required. Bias correction involves simply the addition of a constant to the predictions. This constant is the difference between the average predicted and average measured values for the N samples (i.e., the bias of the predictions). Spiking involves building new PLS models adding those N samples to the calibration dataset.

3. Results and discussion

3.1. Characterization of fruit collected over time from the two orchards

The orchard influenced differently the IQA levels, and the content of the various biochemical analytes measured in the 'Ortanique' juice sampled, although this was dependent on the sampling time/date (Table S1). VJ fruit presented higher weight on 162 d and fruit diameter on 106 d and 162 d, while G fruit presented higher juice content on 106 d. Although SSC and TA were always higher in VJ, MI presented a different pattern, as a result of the variation in the ratio between SSC and TA through time, starting with a higher level in VJ (0 d), and becoming higher in G later on (79 d, 106 d, 172 d). Juice pH exhibited some similarities with MI pattern (higher in VJ on 0 d, and higher in G on 57 d and 79 d) but tended to the same level in both orchards afterwards (106 d and 162 d). In contrast, G tangors exhibited higher firmness and thicker rind, only on 57 d. Differences in these IQA in fruit from orchards located at

various sites have also been reported in 'Newhall' and 'Valencia Late' orange (*Citrus sinensis* (L) Osbeck) (Cavaco et al., 2018, 2021a; Domingues et al., 2021), 'Kinnow' mandarin (Nawaz et al., 2020), and Satsuma mandarin (Li et al., 2021) as a consequence of different soil types, relative humidity, total precipitation and temperature.

Relatively to the biochemical analytes assayed, the content of glucose and fructose were always higher in VJ tangors throughout time (Table S1). Sucrose content was higher in VJ, on 57 d and 106 d. The content of total sugars was higher in VJ through time, except on 79 d and 162 d, which correlates with the higher SSC observed in those days. The content of ascorbic acid was higher in VJ fruit in the beginning (0 d and 57 d), becoming similar in both orchards afterwards. Its contribution for the juice pH is less clear, since the ascorbic acid is only one of the acids contributing to the acidity of the fruit juice. These compounds have also been reported to vary in other citrus varieties when grown under different edaphoclimatic conditions, such as reported for 'Kinnow' mandarin (Nawaz et al., 2020) and 'Satsuma' mandarin (Li et al., 2021).

The various IQA also varied throughout time in each orchard, reflecting their fruit ripening rate (Table S1). Thus, weight and equatorial diameter in VJ orchard, increased over time, being different on 79 d, but no differences were found for fruit weight or size in G orchard. The firmness also decreased over time in both orchards. In the second sampling (57 d), there was a more pronounced decrease, but fruit regained firmness afterwards, although maintained the overall downward trend as ripening evolved.

In G orchard, rind thickness remained constant through time, while in VJ it decreased from 0 d to 57 d, increasing afterwards. Juice pH increased over time in both orchards, following the decreasing acid content over ripening. Consequently, TA decreased through time in both orchards, with constant values from 57 d to 106 d, attaining the minimum on the last sampling date. As expected, SSC increased over time in both orchards, although in a different manner, and started with higher levels in VJ, this difference remaining through time. The maximum level was attained earlier in VJ (106 d), remaining constant afterwards, whereas SSC increased gradually throughout time in G, peaking at 162 d. For MI, an increase was observed over time in both orchards. Results show that juice content decreased over time in both orchards, as fruit ripening progressed.

The evolution of the various IQA along fruit ripening is depending on various factors, namely, the root stock, variety, cultural practices, edaphoclimatic conditions, orchard, tree and fruit site, and tree fruit load (Barman et al., 2015; Cavaco et al., 2021a). Yet, IQA have been shown to follow similar or approximate patterns in 'Newhall' and 'Valencia Late' oranges (Cavaco et al., 2018, 2021a), and in 'Clementine de Corse' (Julhia et al., 2019).

Regarding the biochemical analysis, glucose increased after 79 d in G orchard, remaining constant in VJ over time, except between 57 d and 106 d, when it increased. Fructose increased over time, particularly from 57 d to 79 d, in both orchards. Sucrose increased from 0 d to 57 d, remaining constant until 106 d, and increasing again on the last sampling (162 d) in G orchard. In VJ, sucrose exhibited an increase from 57 d to 79 d, similarly to fructose. Overall, total sugars content increased over time in both orchards. Finally, ascorbic acid decreased from 106 d to 162 d in both orchards, although on 162 d its content attained similar levels to those found at 0 d in VJ (Table S1). As for the other IQA referred above, the temporal variation differences found between orchards regarding these compounds are most probably related to the same or partial factors indicated above (Barman et al., 2015; Cavaco et al., 2021a).

Overall, these samplings resulted in very heterogeneous datasets regarding the various IQA, for the construction of the calibration models and their validation, as presented in Table S2. Briefly, except for firmness, juice pH, sucrose and ascorbic acid content, all the other IQA differed between the various datasets (Table S2). This complies with the importance of including fruit from different origins, seasons and climate

conditions in order to achieve stable and accurate models with an improved prediction performance, through validating calibration models with data from different locations and/or harvest season highlighted by several authors (Nicolai et al., 2007; Wang et al., 2015; Walsh et al., 2020a,b; Cavaco et al., 2021b).

3.2. Spectra

Plots A (VJ) and B (G) in Fig. 1 show the average spectra for each of the five sampling dates. In each sampling day 25 spectra were collected per orchard, but some were rejected because of solar contamination. Thus, the number of spectra considered to calculate these averages varies between 20 and 25. The signal above 1000 nm became increasingly noisy due to the low quantum efficiency of the detector and this is clear even after averaging. Plots C to E illustrate the average spectral variation at three specific wavelengths. Plot C illustrates the average variation in reflectance at 680 nm, essentially determined by the rind chlorophyll content. On the first day, the reflectance is almost null due to the large chlorophyll absorption. But its rind content rapidly decays until the second sampling date. In orchard G there is a reflectance peak in the third sampling day. Its explanation comes from plot D. This plot shows the average variation in reflectance at 840 nm, which is representative of the NIR plateau. In this zone, free of significant absorption bands, the predominant effect is that of scattering, which is determined by the structure and thickness of the rind. The reflectance in the NIR plateau is higher on the first day and then follows a general decreasing trend. However, there is a local peak on the third sampling date, in February, probably due to a change in the rind structure in response to weather conditions. It is this increase in the average level of reflection, which is much more pronounced in the G orchard, that is superimposed across the spectrum, including in the chlorophyll band and in the water band (980 nm). The latter is illustrated in graph E and shows an almost identical pattern to that of the NIR plateau, illustrating well the fact that the absolute level of reflectance in the water band is more determined by the scattering than by the water content itself.

3.3. Calibration models

Calibration models were constructed and validated for all IQA and the analytes presented in Table S2 (SSC, TA, MI, firmness, juice pH, total sugars, sucrose, fructose, glucose and ascorbic acid), following the IV and EV approach, and the various metrics that characterize them are shown in Table 1.

In general, the EV approach resulted in worse models' prediction ability and accuracy for the majority of IQA (Table 1), as expected from previous reports (Cavaco et al., 2018), but the metrics obtained depended on the datasets used for calibration and validation. Thus, while the best results for MI, juice pH, and fructose were obtained for the calibration models constructed with VJ dataset and validated with G orchard data, the opposite was obtained for the remaining IQA and biochemical analytes, for which the best models were constructed with G dataset and validated with VJ (Table 1).

Overall, in EV (G→VJ and VJ→G) MI performed better and more uniformly (SDR consistently ~ 1.5) than the other IQA. Among these, one may distinguish the fair but uniform models (SDR consistently ~ 1.2–1.35 for juice pH, TA, firmness and sucrose content), and the very poor EV results for the remaining IQA (total sugars, fructose, ascorbic acid and, especially, glucose and SSC). The various IQA calibration models will be discussed in the following sections, and MI will be taken as a case study in terms of the calculations and plots shown, for the major IQA used to determine 'Ortanique' harvest date.

As a final note, the performance of the models is related with the X and Y explained variances achieved in calibration. The explained variance in the X block was calculated from $\sum_{i=1}^{nLV} var(T_i) / \sum_{i=1}^N var(T_i)$ where the T_i are the scores of the i -th latent variable, var is the variance

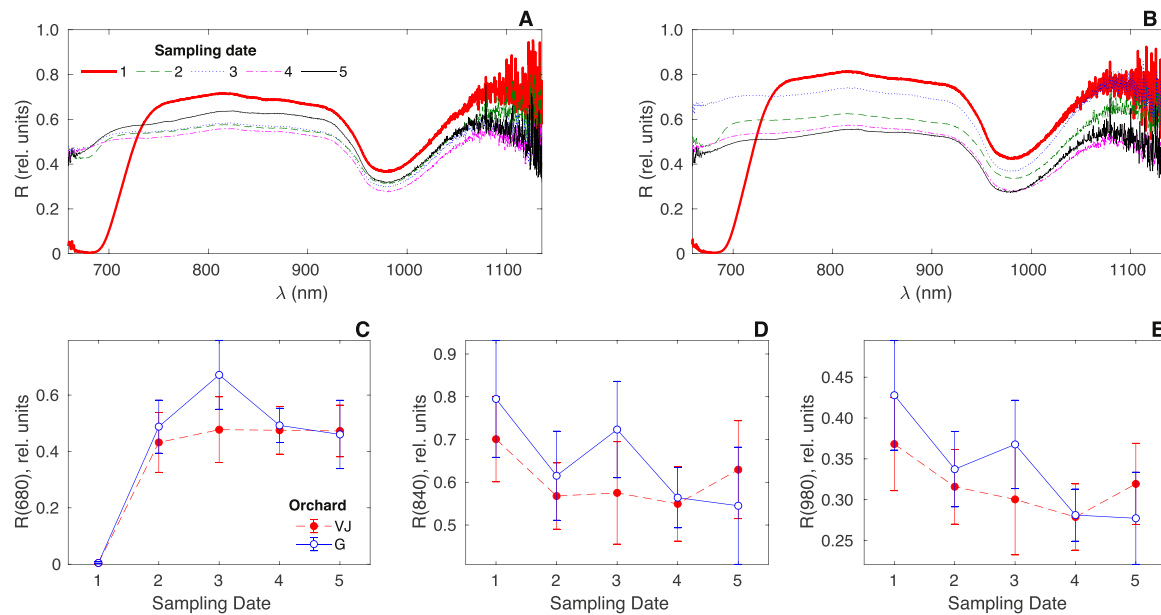


Fig. 1. Average reflectance of the fruit measured in Vale Judeu (VJ) and Gambelas (G). Plots A (VJ) and B (G): average spectra for each of the five sampling dates (20 to 25 spectra averaged in each day). C: variation of the average reflectance at 680 nm. D: variation of the average reflectance at 840 nm. E: variation of the average reflectance at 980 nm.

and N the number of latent variables that cumulatively explain 100 % of the variance. The explained variance in Y was simply read from the R^2 value of the plot $y(\text{predicted})$ vs. $y(\text{measured})$ in the calibration model. But it is the variance explained for block Y in the calibration that determines the model performance, since the variance explained for block X is always high. This can be illustrated by three models with successively worse performance, with $\text{SDR} > 2$, $1 < \text{SDR} < 2$ and $\text{SDR} < 1$, respectively. In the first case, the model for SSC (VJ orchard, $\text{SDR} = 2.27$) explains 98 % of the variance in X and 84 % of the variance in Y ; in the second case, the sucrose model (VJ orchard, $\text{SDR} = 1.38$) explains 82 % and 48 % respectively; and in the latter case, the model for glucose (G orchard, $\text{SDR} = 0.94$) explains 96 % and 39 % respectively.

3.3.1. SSC, TA, firmness and juice pH

As expected, the calibration models obtained for these IQA performed better under IV than under EV conditions, as previously reported for 'Newhall' orange (Cavaco et al., 2018). However, this depended on the IQA and on the dataset used to build and validate the models (Tables S2, 1). The SDR values varied between 1 and 2, with some exceptions reaching $\text{SDR} > 2$. The 20 SDR values shown in Table 1 for SSC, TA, firmness and juice pH, are distributed by these classes as follows: no prediction: 2 models (10 %); poor: 9 (45 %); coarse: 7 (35 %); good: 2 (10 %). Discriminating between IV and EV reveals their unequal performances. The SDR distribution for IV is (0 % no prediction, 30 % poor, 70 % coarse, 0 % good), and for EV is (10 % no prediction, 60 % poor, 0 % coarse, 20 % good). The conservative approach to interpret these results is to say that EV results are generally poor (median = class 2 = poor predictive capacity), failing to prove the applicability of the spectroscopic method in new orchards. The more realistic view, however, is that the IV performance is the limit to which EV will tend if one includes enough variability (i.e., enough number of different orchards) in the calibration set, as previously shown for 'Newhall' orange from different orchards and harvest seasons (Cavaco et al., 2018). The IV median is class 3 = coarse prediction, which means that one may reasonably expect that the spectroscopic method will deliver at least coarse predictions for the IQA in the future.

Table 2 summarizes results obtained by other authors in similar investigations. The metrics obtained in the current study were also inserted for comparison purposes. It refers to calibration models to predict SSC,

TA, firmness and juice pH in mandarins, tangerines and clementines, independently of the preprocessing and chemometrics approaches followed. Both the rind thickness (Wang et al., 2016), its biochemical (Jamshidi et al., 2014), and bulk optical properties of the various citrus fruit (Sun et al., 2020a,b, 2021), have been reported to influence the vis-NIR spectral data, urging to choose the most similar fruit for the comparison. For SSC in IV there are mainly two groups of results: most studies presented $\text{RMSEP} \sim 0.6$ to 0.8 % and $R^2 \sim 0.5$ to 0.9 , where the results obtained in the current study fit; there is also a smaller group of two reports claiming considerably better results, with $\text{RMSEP} \sim 0.3$ to 0.4 % and $R^2 \geq 0.9$, although the low number of samples in Santos et al. (2020) requires a critical regard. The fact that different spectral ranges yield approximately equal results suggests that SSC may be assessed either in the SW-NIR (700–1100 nm) or in the NIR (1000–2500 nm) ranges. Results obtained by the other authors for TA in IV have $\text{RMSEP} \leq 0.18$ % citric acid, while for this work it was obtained a $\text{RMSEP} \sim 0.24$ – 0.27 % citric acid. However, the R^2 values presented here are on par with the best reported by the other authors (Antonucci et al., 2011; Torres et al., 2019). This is probably because the range of TA values in the current study is larger than in those studies, which might have increased the TA variability. On the other side, the results regarding juice pH in IV are uniform, being equivalent to those reported by other authors (Table 2). Finally, the firmness in IV is the most difficult IQA to compare due to differences in the procedure used (compression vs. maximum penetration force, diameter of the compressing plate, etc.). For example, the range of firmness in this work was 64–300 N, while in Sánchez et al. (2013a) it was 2–20 N. To solve this discrepancy, a column with the resolution, $\text{res}(\%)$, as defined in Sec. 2.4.3 was added for firmness, which allows the different models to be compared with this non-dimensional parameter. Recently, Walsh et al. (2020b) reviewed the application of vis-NIR technology in the fruit sector and described the difficulty of assessing firmness non-invasively. Minor changes in pectin are unlikely to be detected by NIRS. Cell wall adherence and shape related to firmness changes might result in changes in light scattering. Although detectable, those changes couldn't be linked to the firmness level observed.

The last lines of Table 2 refer to EV, which use is scarce for mandarins according to the literature. The comparison is performed with the results of Guthrie et al. (2005), specifically with the entry "Location" in their Table 1, since it corresponds to the same situation described in this

Table 2
vis-NIRS models overview (mandarins, tangerines, clementines) found in the literature for comparison purposes with the models obtained in this work. N_{fruit} = number of fruit; N_{orch} = number of orchards; % res = resolution; SSC = soluble solids content; RMSEP = root mean square of prediction; R^2 = regression coefficient; SEP = standard error of prediction; SEP (c) = bias-corrected standard error of Prediction.

INTERNAL VALIDATION (IV)									
SSC (%)	TA (% citric acid)			Juice pH			Firmness (N)		
	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP	R^2
0.75–0.93	0.69–0.79	0.24–0.27	0.69–0.73	0.10–0.19	0.40–0.80	21–40	0.29–0.69	14–18	224
0.77 ⁽¹⁾	0.5	0.15 ⁽¹⁾	0.56	0.11 ⁽¹⁾	0.61	3.0 ⁽¹⁾	0.15	17	234
0.33	0.88	–	–	0.18	0.66	8.5	0.69	20	200
0.58	0.71	0.06	0.66	–	–	11–13 ⁽²⁾	–	–	400
–	–	–	–	–	–	–	0.08–0.28	18	233
0.71–0.84 ⁽²⁾	0.57–0.75	0.13–0.18 ⁽²⁾	0.65–0.76	0.11 ⁽²⁾	0.74	–	–	–	230
0.38 ⁽³⁾	0.96 ⁽³⁾	0.09 ⁽³⁾	0.51 ⁽³⁾	–	–	–	–	–	25 ⁽³⁾
0.40 ⁽⁴⁾	0.97 ⁽⁴⁾	0.10 ⁽⁴⁾	0.92 ⁽⁴⁾	–	–	–	–	–	21 ⁽⁴⁾
0.65	0.86	0.09	0.44	–	–	–	–	–	153

EXTERNAL VALIDATION (EV)									
SSC (%)	TA (% citric acid)			Juice pH			Firmness (N)		
	R^2	RMSEP	R^2	bias	RMSEP	R^2	bias	RMSEP	R^2
1.61–1.83	0.43–0.60	–1.17–1.53	0.69	–0.28–0.30	0.59–0.60	0.19–0.20	–0.09–0.09	29–32	0.34–0.38
0.41–0.53	0.35–0.75	0.31–0.85	–	–	–	–	–	–	0.24–7.8
(1) SEP (c); (2) SEP; (3) clementines; (4) tangerines; (5) fruit acquired in a local market.									

report, which is the validation on another orchard, but in the same season. Overall, our EV models performed worse, with higher RMSEP values. The plausible reason for this result, as explained above, is most likely related to the very high ripening variability in fruit between orchards.

3.3.2. Biochemical analytes

The calibration models obtained for sucrose, total sugars and ascorbic acid exhibited a medium prediction capacity (Table 1). For these compounds, the best calibration model was obtained for the ascorbic acid content followed sequentially by the total sugars and the sucrose content. Calibration models for glucose and fructose contents were very weak with no prediction capacity whatsoever (Table 1). Using the model classification described in Sec. 3.3.1 (no prediction/poor/coarse/good), glucose and fructose stand up as the worse models, with respectively three and two non-acceptable calibration models from the five shown in Table 1. This is in contrast with sucrose calibration model, that had a very uniform record of SDR ~ 1.3 (poor) across all the five models. The R^2 values for sucrose were also much better than the other two sugars. This is probably because sucrose is the most abundant sugar in ‘Ortanique’ (a factor of 2, on average). A partial confirmation of this finding is that the total sugars yielded four poor and one coarse model, slightly better than sucrose, likely because of their larger proportion in the juice. However, the total sugars results were not as consistent as those of sucrose content, showing larger fluctuations across the five models, which may be attributable to more variability in the three sugars mixed data. Magwaza et al. (2014) obtained better calibration models for total sugars, sucrose, fructose and glucose content in the rind of ‘Nules’ Clementine in the spectral range 900–1700 nm: RMSEP = 11.88–38.21 mg/g DW, R^2 = 0.56–0.90 and SDR = 1.11–2.53, in an IV approach. The worse results of this work are most probably related to the low SW-NIR light penetration in the pulp of ‘Ortanique’, as previously reported for various thick rind fruit (Arendse et al., 2018). In turn, for ascorbic acid, Santos et al. (2020) obtained a model for tangerine with R^2 = 0.868 and RMSE = 2.360 mg/100 mL and for clementine with R^2 = 0.767 and RMSE = 3.714 mg/100 mL in the spectral range 900–1700 nm. Xudong et al. (2009) obtained a model for vitamin C with R = 0.81 and RMSEP = 2.7 mg/100 mL in the spectral range 400–1040 nm. Our best IV model for ‘Ortanique’ was constructed with VJ dataset with R^2 = 0.59 and RMSEP = 2.31 mg/100 mL juice. The models for both sugars and particularly acids, appear to clearly benefit from the employment of NIR spectrometers based on InGaAs detectors, working in the range 900–2500 nm (see the review by Cavaco et al., 2021b).

3.3.3. Maturation index

Analysis of the MI model is detailed in this section further the concept of model updating. Fig. 2 shows the plots of predicted vs. measured MI, the predictions being obtained from a random resampling (sec. 2.4.4) implementation of PLS models for the VJ fruit alone (Fig. 2A), for the G fruit alone (Fig. 2B), and for the composed fruit dataset VJ + G (Fig. 2C) in an IV approach.

The spectra were subjected to a simple transformation only: absorbance (log(1/R)) followed by SNV. The range was restricted from 700 to 950 nm because extending the range beyond these limits degraded the models’ performance. This means that the most prominent water peak, at 980 nm, does not seem to play an important role for the MI models, which applies to the remaining IQA. Fig. 2D shows the VIP scores for the VJ + G model, which are the average of the VIP scores obtained in each of the 50 resamplings.

The three models delivered essentially the same performance parameters, except for RMSEP, which was clearly lower for the VJ orchard. However, the SDR values were similar for the wider variation range of MI in the G orchard and, consequently, also in the composed VJ + G dataset. Therefore, the VJ calibration model fitted the data better than the other two models. This is also reflected on the slightly better correlation and slope values. The differences between the models for the two orchards will be better understood in the EV discussion.

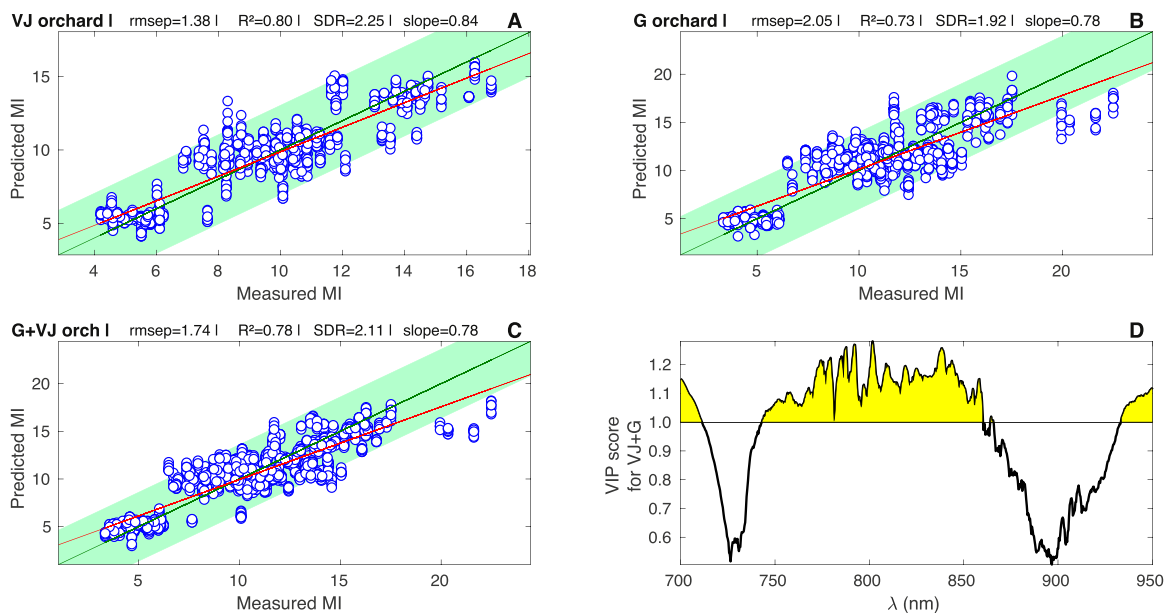


Fig. 2. Internal validation results for the MI calibration models. These results were obtained with random resampling. Plot A: Vale Judeu (VJ) orchard. Plot B: Gambelas (G) orchard. Plot C: global VJ + G model, i.e., with the dataset comprising data from the two orchards. Plot D: VIP scores plot for the calibration model in the aggregate VJ + G. The diagonal band represents one MI standard deviation above and below the central line predicted = measured. RMSEP = root mean square of error of prediction; SDR = standard deviation ratio; R^2 = regression coefficient.

In a previous work, a calibration model for the MI of ‘Newhall’ oranges was developed with the same spectrometer and with the same chemometric analysis (Cavaco et al., 2018). However, the IV ‘Newhall’ results were clearly worse than those obtained here for ‘Ortanique’: SDR ~ 1.4 vs. SDR ~ 2.0 obtained now. This may be related with the results of Sun et al. (2020a,b; 2021), who showed that the bulk optical properties of pulp and rind tissues in citrus have a major impact on the vis-NIRS models. There are significant structural differences regarding the rind thickness and

structural properties of both rind and pulp between ‘Newhall’ orange and ‘Ortanique’, which most probably have a limiting effect on the amount and depth penetration of NIR light and, consequently, on the information acquired from the two citrus varieties. Given the thinner rind in ‘Ortanique’, a deeper light penetration can be expected, resulting in a better assessment of the IQA when compared to ‘Newhall’ orange.

Fig. 3 depicts the results of EV, which are clearly worse than those of IV (Fig. 2, Table 1). In this modality the model is calibrated with data

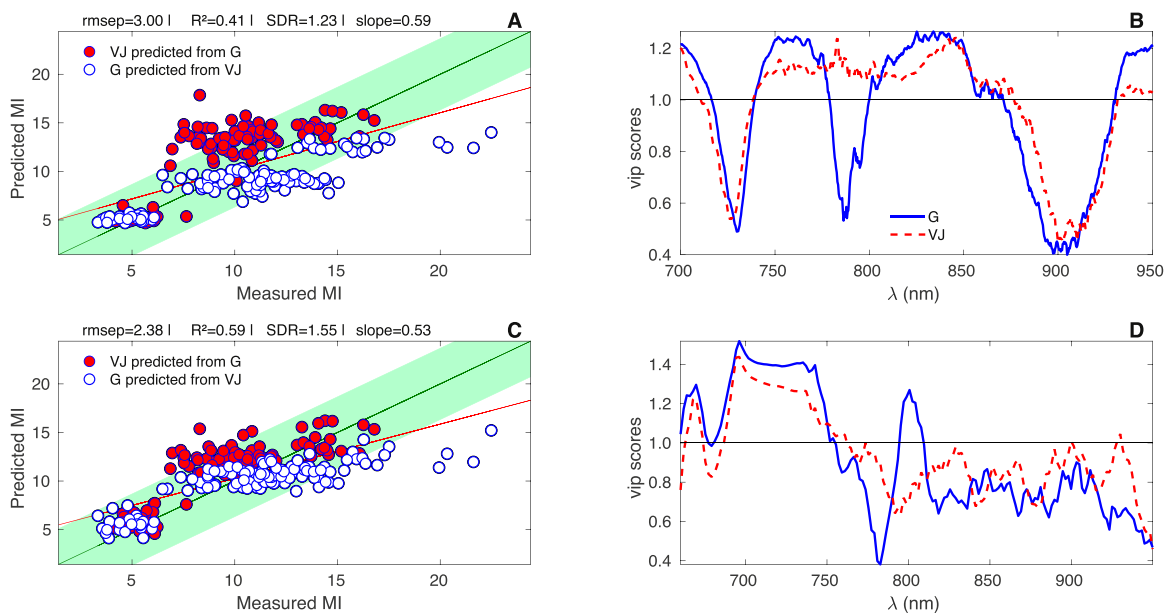


Fig. 3. Plot A: external validation results for the MI models built upon the abs ($\log(1/R)$) + SNV preprocess. The red/filled circles represent the predictions for the VJ orchard from a model calibrated with the G orchard; the white/open circles represent the predictions for the G orchard from a model calibrated in the VJ orchard. The diagonal band represents one MI standard deviation above and below the central line predicted = measured. Plot B: the VIP scores for the model calibrated in the G orchard (blue/continuous line) and for the model calibrated with the VJ orchard (red/dashed line). Plot C: same as plot A, but the preprocess was absorbance + smoothing by Savitzky-Golay (SG) + detrending + derivation by SG (width of SG = 25 points, polynomial orders 0 and 2, respectively for smoothing and derivation). Plot D: same as plot B, but corresponding to the models of plot C. RMSEP = root mean square of error of prediction; SDR = standard deviation ratio; R^2 = regression coefficient.

from one of the orchards and validated with the other. The parameters presented in Fig. 3A, C refer to the validation as a whole, i.e., it is the global statistic for the two validations, G and VJ together (individual performances are listed in Table 1, i.e., for VJ→G and G→VJ separately). In Fig. 3A the filled circles represent the predictions for the VJ orchard from a model calibrated on the G orchard dataset, and vice-versa for the open circles. These models were based on a very simple preprocessing, namely, absorbance + SNV (Table 1).

It is visually clear that the G model overestimates the VJ values and that the VJ model underestimates the G values (Fig. 3A). The origin of this discrepancy may be understood by looking again into the VIP scores, this time on each model separately, which show a strong divergence around 780 nm (Fig. 3B). While this region is the most important for the VJ model (VIP = 1.3 @ 783 nm), it is, on the contrary, essentially devoid of importance in the G model (VIP = 0.5 @ 786 nm). Q (residuals) and T² (Hotelling) plots were further checked (data not shown). Large Q and T² values mean significant variations outside the model and inside the model, respectively. The Q plots showed elevated values around 780 nm for both models. However, the Hotelling T² plot showed much smaller values in the G model compared to those observed in the VJ model. This means that the spectral variations around 780 nm are clearly out-of-the-model for the G dataset but play an ambiguous role in the VJ dataset. It seems that part of that variation, probably irrelevant for a MI model which would have gathered more data, has been captured as relevant by this single season and single orchard model. This illustrates well the limitations of PLS modelling with restricted datasets. The next step was to seek for other transformations that could deliver better external predictions and eliminate the over/under estimations observed in the absorbance + SNV model (Fig. 3A). One such transformation was absorbance + SG + detrending + derivation by SG (width of SG = 25 points, polynomial orders 0 and 2, respectively for smoothing and derivation), in conjunction with a slightly different spectral range of 680–930 nm (Fig. 3C). This transformation gave greater emphasis to the spectral range dominated by the chlorophyll peak and partially suppressed the importance of the other features, particularly the one at 780 nm, that caused the bad EV results. Fig. 3C shows the results of external validation with this transformation, revealing much better results than in Fig. 3A. The corresponding VIP scores plot in Fig. 3D explains the underlying reason for this result: the importance of the models is now condensed in the spectral range 680–750 nm and is relatively consistent for both orchards. This means more homogeneous predictions and a suppression of the over/under estimations found before. However, this transformation yielded slightly worse results in the IV with random resampling validations when compared to the simpler SNV model. This means that it may be used as a patch for EV but it is not a convincing model, since it captures less information.

Summarizing, both external predictions were relatively misplaced, with a bias of approximately ±1 each way. This means that calibration models were based on spectral characteristics that do not reproduce completely when moving from one orchard to another (Guthrie et al., 2005), as illustrated in Fig. 3D. Sánchez et al. (2013a) with ‘Clemevilla’ mandarins, acquired spectra in the same experimental conditions as described for Torres et al. (2019) and obtained a MI model for 188 fruit (116 calibration, 60 validation) with SEP(c) = 1.46, bias = -0.10, R² = 0.53 and slope = 0.85 in the spectral range of 1600–2400 nm, and in the IV mode. This is approximately equivalent to the performance of our best IV model (VJ). However, the other two models (G and VJ + G) have larger RMSEP values and performed worse. Nonetheless it is interesting to note that the SDR, R and slope values are equivalent or even slightly better than those obtained by Sánchez et al. (2013a).

3.3.4. Recalibration of MI models

The search for a better transformation to cope with external validation is an *a posteriori* procedure, in the sense that it is performed to tune the validation. This is obviously “cheating” the calibration procedure and was presented in Fig. 3C for the sole purpose of help to

understand the cause of the poor results observed in EV. Therefore, changing the spectral transformation is not a solution for *real-life* situations. In a *real-life* situation there is a model calibrated with a chosen transformation and the newly acquired spectra will be transformed in the same way. To improve the models’ prediction capacity, it is necessary to introduce a model update in the process. An update model technique allows to “rescue” an inefficient model in order to produce acceptable predictions for a new orchard or batch. It typically requires a small number of recalibration samples, which represents much less effort than building a new model for the new orchard. Fig. 4 shows the results of two such methods, namely bias correction and spiking. The calibration models depicted in Fig. 4 use the absorbance + SNV transformed spectra according to the initial choice as input.

Thus, the previous Fig. 3A may be seen as the no-update result. Bias correction corresponds simply to the addition of a constant to the predictions. This constant is the difference between the mean of the predictions and the mean of the corresponding true values for a given number of recalibration samples (that will be referred hereafter as N_{recal}) taken from the EV dataset. Fig. 4A shows the results obtained with bias correction for $N_{recal} = 20$ taken from the whole dataset, irrespective of the sampling date. This is a typical result, since the 20 recalibration samples were chosen randomly, producing slightly different results at each run. The value of RMSEP = 2.34 was lower than that obtained initially, namely RMSEP = 3.00 (Fig. 3A), and all the other performance parameters improved, e.g., R², from 0.41 to 0.60.

On the other side, the procedure of spiking implies mixing N_{recal} validation samples, also taken from the whole dataset, with the calibration samples to build a new model. Fig. 4C shows the results of spiking for $N_{recal} = 20$. The model performance was improved in comparison to the bias correction, presenting a lower RMSEP (= 1.87) than that calculated for the model after bias correction (RMSEP = 2.34) (Fig. 4A, C). Fig. 4B shows the variation of RMSEP as function of N_{recal} for the case G→VJ. The error bars are obtained from a 20-fold random resampling of the recalibration samples, for each value of N_{recal} . Clearly, the bias correction is more effective in lowering RMSEP when N_{recal} is small (<8) but it stagnates beyond 10 recalibration samples while spiking keeps improving steadily. This is because 10 recalibration samples already offer a good estimate of bias in validation, while adding more recalibration samples always improves the PLS model fit. Fig. 4D shows typical plots of the models’ VIP scores after the inclusion of 20 recalibration samples (note that only spiking changes the initial VIP scores depicted in Fig. 3B). The two curves approached in the band around 780 nm, which was the major source of discrepancy in Fig. 3B. Therefore, spiking brought the G model closer to the VJ model, in terms of wavelength band relevance, with the consequent improvement in the EV results (Figs. 3A, B, 5C, D).

The final piece of analysis of the recalibration process is depicted in Fig. 5.

In Fig. 4A and C, both the bias correction and the spiking procedures were calculated from random choices of validation samples. However, this implies the assumption that all the data is available, that is, this approach is strictly valid only at the end of the season. But the objective of this study was to follow-up the ripening of the orchards, and this implies a series of samplings dates spaced in time. Therefore, model updating in the first visit to the orchard can be performed only with fruit from that date; and model updating at the second visit can only rely on fruit from the first and second visits, and so on. In the case of spiking, the most efficient way to perform the model update is retroactively. This means that the recalibration samples taken at date N are used together with those collected on dates 1, 2, ..., N-1 to update the model. The updated model is then used to make predictions on the validation samples from date 1 to date N. In the case of bias correction, the most efficient way is to recalibrate daily. This means that the recalibration samples taken at date N are used to correct the predictions for that date only. Fig. 5A and B simulate this scenario of orchard follow-up. The x axis contains the number of the visits to the orchard (sampling date), which

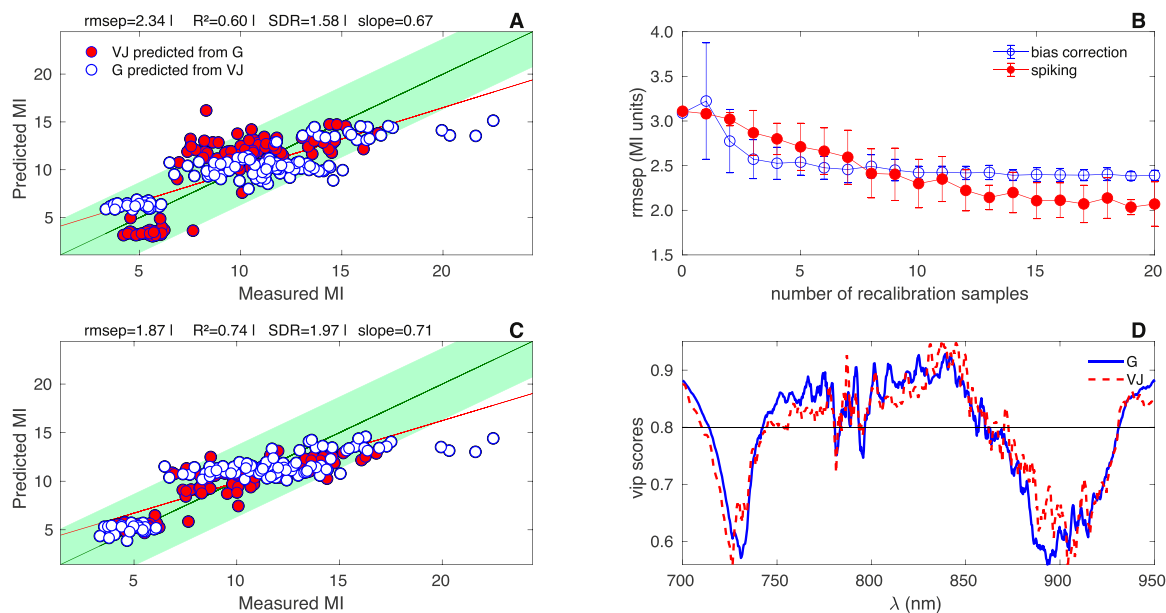


Fig. 4. External validation results for the MI models built upon the $\text{abs}(\log(1/R)) + \text{SNV}$ preprocess and using either bias correction or spiking. Plot A: external validation results with bias correction. The red/filled circles represent the predictions for the VJ orchard, corrected for bias, from a model calibrated in the G orchard; the white/open circles represent the predictions for the G orchard, corrected for bias, from a model calibrated in the VJ orchard. The diagonal band represents one MI standard deviation above and below the central line predicted = measured. Plot B: variation of RMSEP as function of the number of recalibration samples for the case G→VJ. The bias correction values are the blue/open circles and the spiking values are the red/filled circles. Plot C: external validation results with spiking. The red/filled circles represent the predictions for the VJ orchard, from a model calibrated with a dataset constituted by the G orchard fruit plus 20 random fruit from the VJ orchard; the white/open circles represent the predictions for the G orchard, from a model calibrated with a dataset constituted by the VJ orchard fruit plus 20 random fruit from the G orchard. Plot D: VIP scores corresponding to the spiking models of plot C. The model calibrated with all G fruit plus 20 V J fruit in blue/continuous line and the model calibrated with all VJ fruit plus 20 G fruit in red/dashed line. RMSEP = root mean square of prediction; SDR = standard deviation ratio; R² = regression coefficient.

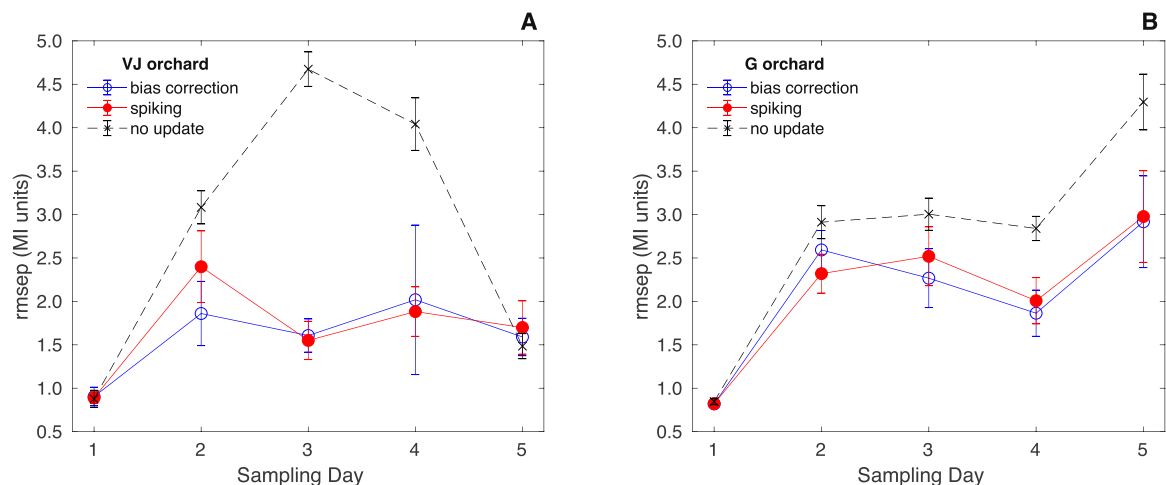


Fig. 5. Plot A: RMSEP as a function of the sampling date, for the predictions in VJ orchard. Black/dashed curve: no model updating. Red/filled symbols: spiking results, with 5 spiking samples per date. The recalibration samples taken at day N are used together with those collected on days 1,2,...,N-1 to update the model. Blue/open symbols: bias correction results, with 5 recalibration samples per date. In this case the recalibration samples taken at day N are used to correct the predictions of that day only. Plot B: same as plot B, but for the G orchard. RMSEP = root mean square of prediction.

was five in this work. The y axis shows the RMSEP values after the recalibration with five samples at each date. Fig. 5C depicts the data obtained from the combination G→VJ and vice-versa for Fig. 5D. Again, the error bars were obtained from a 20-fold random resampling of the recalibration samples, for each sampling date. The most important conclusions to extract from these results are: i) model updates generally represent an improvement; ii) bias correction and spiking yield approximately the same results; iii) the performance improvement obtained

through model updating was variable. It decreased the RMSEP up to 66 % in G→VJ, but only up to 25 % in VJ → G.

Spiking is a technique used mainly in the context of PLS models for soil analysis (Guerrero et al., 2010). It has also been applied in fruit, such as pears (Mishra et al., 2021) and citrus. Specifically for citrus, it has been used by Magwaza et al. (2014) to improve the prediction of dry matter (DM) in 'Nules' Clementine that the prediction error for dry matter could be decreased by 50 % by the inclusion of four validation

samples into the calibration set. These four samples were chosen on the lower and upper limits of dry matter range. But there is an important difference regarding the current work: all the fruit in the study of Magwaza et al. (2014) were collected at the same time, at commercial maturity, while our fruit were collected over six months of the growing season, which turns spiking a more difficult task. Mishra et al. (2021) investigated the improvement of EV predictions by using variable selection techniques and/or spiking. Although the best results were obtained combining both, it was also shown that spiking alone greatly improved the RMSEP values, in line with the results of the present report.

4. Conclusions

This study demonstrates the strong potential for the use of SW-NIRS as a smart tool to provide an extensive, fast, and real-time assessment of the ripeness of 'Ortanique' citrus fruit on tree. The PLS models provided good, coarse and poor predictions for several IQA in this variety of Algarve Citrus 'with a high impact in the market. Specifically, in internal validation the results are good for MI (SDR = 1.92–2.25), TA (SDR = 1.80–1.94), SSC (SDR = 1.79–2.27), coarse for firmness (SDR = 1.39–1.85), juice pH (SDR = 1.29–2.23) and poor for sucrose (SDR = 1.23–1.38), total sugars (SDR = 1.19–1.62), glucose (SDR = 0.94–1.50), fructose (SDR = 1.03–1.28) and ascorbic acid (SDR = 1.25–1.62). These results degraded in external validation. However, it is expectable that the external validation results tend to those of internal validation by including more fruit lots from other harvest seasons and other orchards.

Simple model updating strategies were tested for external validation of MI, either in the mode of bias correction or spiking. They clearly improved the predictions when the datasets were given as a whole, i.e., without considering the acquisition timeline of fruit spectra. The SDR values improved from SDR = 1.00–1.39 (poor) to coarse/good values of SDR = 1.87–2.34.

The efficiency of model updating in the context of orchard follow-up was also simulated. The simulations showed a decrease of RMSEP (measured in each sampling date) relatively to the no-update results, in the range of 25 % to 66 %. This level of improvement was observed as early as in the second sampling date.

Funding sources

This work was supported by FCT - Fundação para a Ciência e a Tecnologia, Portugal (post-doc fellowship SFRH/BPD/101634/2014), the CEOT strategic projects UID/Multi/00631/2013 and UID/Multi/00631/2019 and NIBAP: Núcleo de Investigação em Biotecnologia e Agricultura de Precisão | ALG-01-0247-FEDER-037303. Ana M. Cavaco is presently funded by Portuguese national funds from FCT - Fundação para a Ciência e a Tecnologia, Portugal in the ambit of the contract program DL 57/2016/CP1361/CT0031.

CRedit authorship contribution statement

Rosa Pires: Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing. **Rui Guerra:** Conceptualization, Methodology, Software, Validation, Formal analysis, Data curation, Writing - original draft, Writing - review & editing, Supervision, Funding acquisition. **Sandra P. Cruz:** Formal analysis, Investigation, Writing - review & editing. **M. Dulce Antunes:** Methodology, Writing - review & editing, Supervision, Funding acquisition. **António Brázio:** Investigation. **Andreia M. Afonso:** Investigation. **Mariana Daniel:** Investigation. **Thomas Panagopoulos:** Methodology, Writing - review & editing, Funding acquisition. **Isabel Gonçalves:** Resources. **Ana M. Cavaco:** Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors acknowledge CACIAL (Almancil, Portugal) for providing the access to the Citrus orchards followed in this study. Márcia Rosendo Lucas, from CACIAL, is also acknowledged for the supervision of orchards management.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.postharvbio.2021.111750>.

References

- Antonucci, F., Pallottino, F., Paglia, G., Palma, A., D'Aquino, S., Menesatti, P., 2011. Non-destructive estimation of mandarin maturity status through portable VIS-NIR spectrophotometer. *Food Bioproc. Tech.* 4, 809–813. <https://doi.org/10.1007/s11947-010-0414-5>.
- 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. *J. Food Eng.* 217, 11–23. <https://doi.org/10.1016/j.jfoodeng.2017.08.009>.
- Atlas do Ambiente, 1995. Agência Portuguesa Do Ambiente, Scale 1:1.000.000. Lisbon, Portugal <https://sniamb.apambiente.pt/content/geo-visualizador?language=pt-pt> (Accessed 14-10-2020).
- Barman, K., Ahmad, M.S., Siddiqui, M.W., 2015. Factors affecting the quality of fruit and vegetables: recent understandings. In: Siddiqui, M.W. (Ed.), *Postharvest Biology and Technology of Horticultural Crops*. Apple Academic Press (Taylor and Francis Group), Oakville, pp. 1–50.
- Barnes, R.J., Dhanoa, M.S., Lister, S.J., 1989. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Appl. Spectrosc.* 43 (5), 772–777.
- Bizzani, M., Flores, D.W.M., Colnago, L.A., Ferreira, M.D., 2017. Non-invasive spectroscopic methods to estimate orange firmness, peel thickness, and total pectin content. *Microchem. J.* 133, 168–174. <https://doi.org/10.1016/j.microc.2017.03.039>.
- Borba, K.R., Spricigo, P.C., Aykas, D.P., Mitsuyuki, M.L., Colnago, L.A., Ferreira, M.D., 2021. Non-invasive quantification of vitamin C, citric acid, and sugar in 'Valencia' oranges using infrared spectroscopies. *J. Food Sci. Technol.* 58, 731–738. <https://doi.org/10.1007/s13197-020-04589-x>.
- Canavari, M., Centonze, R., Hingley, M., Spadoni, R., 2010. Traceability as part of competitive strategy in the fruit supply chain. *Br. Food J.* 112, 171–186. <https://doi.org/10.1108/00070701011018851>.
- Cavaco, A.M., Pires, R., Antunes, M.D., Panagopoulos, T., Brázio, A., Afonso, A.M., Silva, L., Lucas, M.R., Cadeiras, B., Cruz, S.P., Guerra, R., 2018. Validation of short wave near infrared calibration models for the quality and ripening of 'Newhall' orange on tree across years and orchards. *Postharvest Biol. Technol.* 141, 86–97. <https://doi.org/10.1016/j.postharvbio.2018.03.013>.
- Cavaco, A.M., Cruz, S.P., Antunes, M.D., Guerra, R., Pires, R., Afonso, A.M., Brázio, A., Silva, L., Lucas, M.R., Daniel, M., Panagopoulos, T., 2021a. Spatiotemporal modelling of the quality and ripening of two cultivars of "Algarve Citrus" orchards at different edaphoclimatic conditions. *Postharvest Biol. Technol.* 172, 111386 <https://doi.org/10.1016/j.postharvbio.2020.111386>.
- Cavaco, A.M., Passos, D., Pires, R.M., Antunes, M.D., Guerra, R., 2021b. Nondestructive Assessment of Citrus Fruit Quality and Ripening by Visible-Near Infrared Reflectance Spectroscopy [Online First]. *IntechOpen*. <https://doi.org/10.5772/intechopen.95970>. Available from: <https://www.intechopen.com/online-first/non-destructive-assessment-of-citrus-fruit-quality-and-ripening-by-visible-near-infrared-reflectance>.
- Clark, R.N., Roush, T.L., 1984. Reflectance spectroscopy: quantitative analysis techniques for remote sensing applications. *J. Geophys. Res. Solid Earth* 89 (B7), 6329–6340.
- de Boor, C., 1980. A practical guide to splines. *Math. Comput.* 34 (149), 325–326. <https://doi.org/10.2307/2006241>.
- Domingues, A.R., Marcolini, C.D.M., Gonçalves, C.H.D.S., Gonçalves, L.S.A., Roberto, S. R., Carlos, E.F., 2021. Fruit ripening development of 'Valencia' orange trees grafted on different 'Trifoliata' hybrid rootstocks. *Horticulturae* 7, 3. <https://doi.org/10.3390/horticulturae7010003>.
- Golic, M., Walsh, K., Lawson, P., 2003. Short-wavelength near-infrared spectra of sucrose, glucose, and fructose with respect to sugar concentration and temperature. *Appl. Spectrosc.* 57, 139–145. <https://doi.org/10.1366/000370203321535033>.
- Gómez, A.H., He, Y., Pereira, A.G., 2006. Non-destructive measurement of acidity, soluble solids and firmness of Satsuma mandarin using Vis/NIR-spectroscopy techniques. *J. Food Eng.* 77, 313–319. <https://doi.org/10.1016/j.jfoodeng.2005.06.036>.

- Guerrero, C., Zornoza, R., Gómez, I., Mataix-Beneyto, J., 2010. Spiking of NIR regional models using samples from target sites: effect of model size on prediction accuracy. *Geoderma* 158, 66–67. <https://doi.org/10.1016/j.geoderma.2009.12.021>.
- Guthrie, J.A., Reid, D.J., Walsh, K.B., 2005. Assessment of internal quality attributes of mandarin fruit. 2. NIR calibration model robustness. *Aust. J. Agric. Res.* 56, 417–426. <https://doi.org/10.1071/AR04299>.
- Haleem, A., Khan, S., Khan, M.I., 2019. Traceability implementation in food supply chain: a grey-DEMATEL approach. *Inf. Process. Agric.* 6, 335–348. <https://doi.org/10.1016/j.inpa.2019.01.003>.
- Hopkins, D., 2001. What is a Norris derivative? *Nir News* 12 (1), 3. <https://doi.org/10.1255/nirn.611>.
- Jamshidi, B., Minaei, S., Mohajerani, E., Ghasseman, H., 2014. Prediction of soluble solids in oranges using visible/near-infrared spectroscopy: effect of peel. *Int. J. Food Prop.* 17 (7), 1460–1468. <https://doi.org/10.1080/10942912.2012.717332>.
- Julhía, L., Belmin, R., Meynard, J.-M., Pailly, O., Casabianca, F., 2019. Acidity drop and chloriation in Clementine: implications for fruit quality and harvesting practices. *Front. Plant Sci.* 10, 754. <https://doi.org/10.3389/fpls.2019.00754>.
- Ladaniya, M.S., 2008. *Citrus fruit: Biology, Technology and Evaluation*. Academic Press (Elsevier), San Diego, CA, USA, p. 558. ISBN: 978-0-12-374130-1.
- Lado, J., Rodrigo, M.J., Zacarias, L., 2014. Maturity indicators and citrus fruit quality. *Stewart Postharvest Rev.* 2, 1–6.
- Li, Z., Jin, R., Yang, Z., Wang, X., You, G., Guo, J., Zhang, Y., Liu, F., Pan, S., 2021. Comparative study on physicochemical, nutritional and enzymatic properties of two Satsuma mandarin (*Citrus unshiu* Marc.) varieties from different regions. *J. Food Compos. Anal.* 95, 103614. <https://doi.org/10.1016/j.jfca.2020.103614>.
- Liu, C., Yang, S.X., Deng, L., 2015. A comparative study for least angle regression on NIR spectra analysis to determine internal qualities of navel oranges. *Expert Syst. Appl.* 42, 8497–8503. <https://doi.org/10.1016/j.eswa.2015.07.005>.
- Magwaza, L.S., Opara, U.L., Nieuwoudt, H., Cronje, P.J.R., Saeys, W., Nicolai, B., 2012. NIR spectroscopy applications for internal and external quality analysis of citrus fruit—a review. *Food Bioproc. Tech.* 5, 425–444. <https://doi.org/10.1007/s11947-011-0697-1>.
- Magwaza, L.S., Opara, U.L., Cronje, P.J.R., Landahl, S., Nieuwoudt, H.H., Mouazen, A. M., Nicolai, B.M., Terry, L.A., 2014. Assessment of rind quality of ‘Nules Clementine’ mandarin fruit during postharvest storage: 2. Robust Vis/NIRS PLS models for prediction of physico-chemical attributes. *Sci. Hortic.* 165, 421–432. <https://doi.org/10.1016/j.scienta.2013.09.050>.
- Martens, H., Jensen, S.A., Geladi, P., 1983. Multivariate linearity transformations for near infrared reflectance spectroscopy. In: Christie, O.H.J. (Ed.), *Proc. Nordic Symp. Applied Statistics*. Stokkland Forlag, Stavanger, Norway, pp. 205–234.
- Mishra, P., Woltering, E., Brouwer, B., Hogeveen-van Echtelt, E., 2021. Improving moisture and soluble solids content prediction in pear fruit using near-infrared spectroscopy with variable selection and model updating approach. *Postharvest Biol. Technol.* 171, 111348. <https://doi.org/10.1016/j.postharvbio.2020.111348>.
- Nawaz, R., Abbasi, N.A., Hafiz, I.A., Khalid, A., 2020. Impact of climate variables on fruit internal quality of Kinnow mandarin (*Citrus nobilis* Lour x *Citrus deliciosa* Tenora) in ripening phase grown under varying environmental conditions. *Sci. Hortic.* 265, 109235. <https://doi.org/10.1016/j.scienta.2020.109235>.
- Ncama, K., Opara, U.L., Tesfay, S.Z., Fawole, O.A., Magwaza, L.S., 2017. Application of Vis/NIR spectroscopy for predicting sweetness and flavour parameters of ‘Valencia’ orange (*Citrus sinensis*) and ‘Star Ruby’ grapefruit (*Citrus x paradisi* Macfad). *J. Food Eng.* 193, 86–94. <https://doi.org/10.1016/j.jfoodeng.2016.08.015>.
- Ncama, K., Magwaza, L.S., Tesfay, S.Z., Mditshwa, A., Mbili, N.C., 2020. In-field application of portable NIR to assess ‘Valencia’ orange fruit maturity. *Acta Hortic.* 1275, 61–68. <https://doi.org/10.17660/ActaHortic.2020.1275.9>.
- Nicolai, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K.L., Lammertyn, J., 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review. *Postharvest Biol. Technol.* 46, 99–118. <https://doi.org/10.1016/j.postharvbio.2007.06.024>.
- Pasquini, C., 2018. Near infrared spectroscopy: a mature analytical technique with new perspectives – a review. *Analitica Chimica Acta* 1026, 8–36. <https://doi.org/10.1016/j.aca.2018.04.004>.
- Peirs, A., Scheerlinck, N., Nicolai, B.M., 2003. Temperature compensation for near infrared reflectance measurement of apple fruit soluble solids contents. *Postharvest Biol. Technol.* 30 (3), 233–248. [https://doi.org/10.1016/S0925-5214\(03\)00118-2](https://doi.org/10.1016/S0925-5214(03)00118-2).
- Pires, R., Cavaco, A.M., Antunes, M.D., Panagopoulos, T., Brázio, A., Afonso, A.M., Daniel, M., Rosendo, M.L., Gonçalves, I., Guerra, R., 2019. Ripening forecast of ‘Ortanique’ (*Citrus reticulata* Blanco x *Citrus sinensis* (L) Osbeck) based on georeferenced internal quality parameters. In: XXIII Meeting of the Spanish Society of Plant Physiology and the XVI Hispano-Portuguese Congress of Plant Physiology. 26–28 June 2019, Pamplona, Spain.
- Rinnan, Å., Andersson, M., Ridder, C., Engelsen, S.B., 2014. Recursive weighted partial least squares (rPLS): an efficient variable selection method using PLS. *J. Chemom.* 28 (5), 439–447. <https://doi.org/10.1002/cem.2582>.
- Sánchez, M.T., De la Haba, M.J., Pérez-Marín, D., 2013a. Internal and external quality assessment of mandarins on-tree and at harvest using a portable NIR spectrophotometer. *Comput. Electron. Agric.* 92, 66–74. <https://doi.org/10.1016/j.compag.2013.01.004>.
- Sánchez, M.T., De la Haba, M.J., Serrano, I., Pérez-Marín, D., 2013b. Application of NIRS for nondestructive measurement of quality parameters in intact oranges during on-tree ripening and at harvest. *Food Anal. Methods* 6, 826–837. <https://doi.org/10.1007/s12161-012-9490-7>.
- Santos, C.S.P., Cruz, R., Gonçalves, D.B., Queirós, R., Bloore, M., Kovács, Z., Hoffmann, I., Casal, S., 2020. Non-destructive measurement of the internal quality of citrus fruit using a portable NIR device. *J. AOAC Int.*, qsa115. <https://doi.org/10.1093/jaoacint/qsaa115>.
- Savitzky, A., Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* 36 (8), 1627–1639. <https://doi.org/10.1021/ac60214a047>.
- Singh, K.K., Reddy, B.S., 2006. Post-harvest physico-mechanical properties of orange peel and fruit. *J. Food Eng.* 73, 112–120. <https://doi.org/10.1016/j.jfoodeng.2005.01.010>.
- Stivastava, S., Sadistap, S., 2018. Non-destructive sensing methods for quality assessment of on-tree fruit: a review. *J. Food Meas. Charact.* 12, 497–526. <https://doi.org/10.1007/s11694-017-9663-6>.
- Sun, C., Van Beers, R., Aernouts, B., Saeys, W., 2020a. Bulk optical properties of citrus tissues and the relationship with quality properties. *Postharvest Biol. Technol.* 163, 111127. <https://doi.org/10.1016/j.postharvbio.2020.111127>.
- Sun, C., Aernouts, B., Saeys, W., 2020b. Effects of harvest time, fruit size and cultivar on the bulk optical properties of Satsuma mandarin. *Postharvest Biol. Technol.* 111412. <https://doi.org/10.1016/j.postharvbio.2020.111412>.
- Sun, C., Aernouts, B., Van Beers, R., Saeys, W., 2021. Simulation of light propagation in citrus fruit using Monte Carlo multi-layered (MCML) method. *Postharvest Biol. Technol.* 291, 110225. <https://doi.org/10.1016/j.jfoodeng.2020.110225>.
- Torres, I., Pérez-Marín, D., De la Haba, M.J., Sánchez, M.T., 2017. Developing universal models for the prediction of physical quality in citrus fruit analysed on-tree using portable NIRS sensors. *Biosyst. Eng.* 153, 140–148. <https://doi.org/10.1016/j.biosystemseng.2016.11.007>.
- Torres, I., Sánchez, M.T., De la Haba, M.J., Pérez-Marín, D., 2019. LOCAL regression applied to a citrus multispecies library to assess chemical quality parameters using near infrared spectroscopy. *Spectrochim. Acta A. Mol. Biomol. Spectrosc.* 217, 206–214. <https://doi.org/10.1016/j.saa.2019.03.090>.
- Uniprofrutal, 2003. *Manual do Utilizador, IGP “Citros do Algarve” Faro, Uniprofrutal*. APAGAL, p. 12.
- Walsh, K.B., McGlone, V.A., Han, D.C., 2020a. The uses of near infra-red spectroscopy in postharvest decision support: a review. *Postharvest Biol. Technol.* 163, 111139. <https://doi.org/10.1016/j.postharvbio.2020.111139>.
- Walsh, K.B., Blasco, J., Zude-Sasse, M., Sun, X., 2020b. Visible-NIR ‘point’ spectroscopy in postharvest fruit and vegetable assessment: the science behind three decades of commercial use. *Postharvest Biol. Technol.* 168, 111246. <https://doi.org/10.1016/j.postharvbio.2020.111246>.
- Wang, H., Peng, J., Xie, C., Bao, Y., He, Y., 2015. Fruit quality evaluation using spectroscopy technology: a review. *Sensors* 15, 11889–11927. <https://doi.org/10.3390/s150511889>.
- Wang, A., Wang, C., Xie, L., 2016. Influence of the peel on predicting soluble solids content of navel oranges using visible and near-infrared spectroscopy. *Trans. ASABE* 59 (1), 31–37. <https://doi.org/10.13031/trans.59.11152>.
- Xudong, S., Hailiang, Z., Yande, L., 2009. Nondestructive assessment of quality of Nanfeng mandarin fruit by a portable near infrared spectroscopy. *Int. J. Agric. Biol. Eng.* 2 (1), 65–71.
- Zude, M., Pflanz, M., Kaprielian, C., Aivazian, B., 2008. NIRS as a tool for precision horticulture in citrus industry. *Biosyst. Eng.* 99, 455–459. <https://doi.org/10.1016/j.biosystemseng.2007.10.016>.