Revista Ciência Agronômica, v. 52, n. 1, e20196949, 2021
Centro de Ciências Agrárias - Universidade Federal do Ceará, Fortaleza, CE
www.ccarevista.ufc.br

*Scientific Article*
ISSN 1806-6690

# Monitoring vitamin C extraction using multivariate calibration models by NIR[1]

## Monitoramento da extração de vitamina C empregando modelos de calibração multivariados por espectroscopia NIR

**Luciana Maria Herculano da Silva[2], Lívia Paulia Dias Ribeiro[3], Bianca Carvalho Costa[2], Ebenezer Oliveira Silva[4] and Maria Raquel Alcântara de Miranda[2]***

**ABSTRACT -** Due to high of vitamin C content, acerola is exploited as source of this vitamin for the enrichment of industrial products. This work aimed to develop a method for monitoring vitamin C content using near infrared (NIR) during extraction procedure from acerola, thereby different processing steps were evaluated. The calibration and validation models were obtained by partial least squares regression with correlation between values by the reference method, spectrophotometry at visible 525 nm, and absorption data by near infrared spectroscopy, 800 to 2500 nm. The most robust quantification model was determined using coefficient of determination ($R^2$), root mean square error of calibration (RMSECV) and root mean square error of prediction (RMSEP). Vitamin C content ranged from 1,188.39 to 9,959.74 mg. 100 $g^{-1}$, throughout extraction procedure. The obtained RMSEP, 166.27 mg 100 $g^{-1}$, indicates NIR spectroscopy as a promising tool for quantification of vitamin C during extraction from acerola, with the possibility of verifying the content in intermediate stages of production line and moreover, enabling adjustments for correction.

**Key words:** NIR. Ascorbic acid. Extract. Acerola.

**RESUMO -** Devido seu elevado conteúdo de vitamina C, a acerola é explorada como fonte dessa vitamina para o enriquecimento de produtos industriais. Esse trabalho teve como objetivo desenvolver um método de monitoramento do conteúdo de vitamina C usando o infravermelho próximo (NIR), assim foram estudadas diferentes etapas durante sua extração de acerola. Os modelos de calibração e validação foram obtidos por regressão dos quadrados mínimos parciais com a correlação dos valores do método de referência por espectrofotometria no visível a 525 nm e dos dados de absorção por espectroscopia no infravermelho próximo de 800 a 2.500 nm. O modelo de quantificação mais robusto foi determinado utilizando o coeficiente de determinação ($R^2$) e os erros de validação interna (RMSECV) e de previsão (RMSEP). O conteúdo de vitamina C variou de 1.188,39 a 9.959,74 mg. 100 $g^{-1}$, durante todo processamento. O erro de previsão RMSEP, 166,27 mg 100 $g^{-1}$, obtido corrobora com o uso de espectroscopia NIR como uma metodologia promissora para a quantificação durante a extração de vitamina C de acerola, com a possibilidade de verificar o conteúdo em etapas intermediárias da própria linha de produção e permitindo que ajustes sejam feitos para correção.

**Palavras-chave:** NIR. Ácido Ascórbico. Extrato. Acerola.

DOI: 10.5935/1806-6690.20210008
Editor do artigo: Professor Alek Dutra - alekdutra@ufc.br
*Author for correspondence
Received for publication 08/10/2019; approved on 15/09/2020
[1]Parte da Tese do primeiro autor, apresentada ao Curso de Pós-Graduação em Agronomia/Fitotecnia, Universidade Federal do Ceará/UFC
[2]Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará/UFC, *Campus* do Pici, Av. Mister Hull 2297, Fortaleza-CE, Brasil, 60.455-760, luherculano@hotmail.com (ORCID ID 0000-0001-8990-1099), bibi.costa01@hotmail.com (ORCID ID 0000-0002-7461-2638), rmiranda@ufc.br (ORCID ID 0000-0003-0069-7242)
[3]Instituto de Ciências Exatas e da Natureza, Universidade da Integração Internacional da Lusofonia Afro-Brasileira/UNILAB, Rod. CE 060 Km 51. Acarape-CE, Brasil, liviapaulia@unilab.edu.br (ORCID ID 0000-0003-3067-1908)
[4]Embrapa Agroindústria Tropical/EMBRAPA, R. Dra. Sara Mesquita, Fortaleza-CE, Brasil, 60.511-110, ebenezer.silva@embrapa.br (ORCID ID 0000-0002-7396-6637)

## INTRODUCTION

Acerola is well known for its high vitamin C, phenolics and carotenoids content, however and due to its short postharvest life, fruit processing represents a viable alternative to preserve its relevant nutritional properties (GARCIA; SILVA; SEIXAS, 2013; TOLEDO-MARTÍN *et al*., 2016). Moreover, immature acerola has been employed for industrial extraction of vitamin C or ascorbic acid, which may be added to several food matrices. Therefore, countries as United States, Japan and in Europe have shown an 8.5% annual increase in vitamin C enrichment processes, while the main manufacturing companies are: Green Labs LLC, Nutrilite™ from Amway, Naturex, Nature's Power Nutraceuticals Corp., Florida Food Inc., Diana Naturals and Vita Forte. Among these, Nutrilite™ uses acerola grown in organic farms in northeastern region of Brazil (BELWAL *et al*., 2018), which is responsible for almost 70% of the national acerola production where Pernambuco, with 23.11% and Ceará, with 14.32%, standout as the main growers (CALGARO; BRAGA, 2012).

Near infrared spectroscopy (NIR) technology has been employed to determine qualitative and quantitatively organic chemical compounds present in agricultural products and food once it is sensitive to deformations of CH, OH, NH and SH linkages (ARENDSE *et al*., 2017, 2018; CASALE *et al*., 2016; GARCIA; SILVA; SEIXAS, 2013; MAGWAZA *et al*., 2012; TIERNO *et al*., 2016; VENDRAMINI; TRUGO, 2000). Thus, NIR has been advantageously used to monitor quality of food products based on the fact that each sample analyzed has a spectrum that consists of a unique and characteristic pattern (AMODIO *et al*., 2017). As it enables the evaluation of either individual or complex components, NIR becomes an ideal technology to monitor and authenticate complex food matrices (SNYDER *et al*., 2014), however, it is limited by the compound detection threshold of at least 0.1% of the analyzed sample composition (GARCIA; SILVA; SEIXAS, 2013).

NIR together with chemometrics have been successfully used in industrial production lines to assess quality attributes as moisture, pH and acidity, in addition to measuring sugar, soluble solid and vitamin contents in processed pulps of guava and passion fruit (ALAMAR *et al*., 2016). Garcia, Silva and Seixas (2013) used NIR technology to evaluate vitamin C content in a simple extract resultant from pressing acerola and obtained a reliable calibration curve. Arendse *et al*. (2018), used NIR to assess quality of whole pomegranates. Camarês *et al*. (2017), used NIR to control quality by assessing acidity and pH, in addition to sugar, soluble solid and ascorbic acid content in cashew apple and guava nectars. NIR was also used to analyze moisture and soluble sugars of Chinese chestnuts (*Castanea mollíssima* Blume cv. Zaofeng) and results showed this technology may be applied to other nuts (HU *et al*., 2017).

Alamar *et al.* (2016), reported that NIR technology represents an alternative to other quality control approaches that generate chemical residue, require long-lasting procedures or need specialized workmanship. In this context, and since Garcia, Silva and Seixas (2013) used NIR to detect vitamin C from pressed acerola extract, this work aims at using NIR spectroscopy to monitor vitamin C content during the process of its industrial extraction from immature acerola.

## MATERIAL AND METHODS

### Sampling

Immature, green-colored acerola harvested at maximum size produced in organically cultivation system at the state of Ceará, Brazil, were submitted to industrial processing system at Nutrilite™ from Amway do Brasil company located in Tianguá-CE for extraction of concentrated vitamin C. Fruits were selected, deseeded, and processed in subsequent steps to obtain the concentrated vitamin C extract as the following: 1- before the enzymatic treatment with pectin methyl esterase (PME) (step EAE); 2 - after the 30 min enzymatic treatment with PME (step EPE); 3 - after 60 min of calcium hydroxide (CaOH2) addition (step EAH); 4 - after decanting for 3 h and 30 min (step DEC); 5 - during ultrafiltration for 20 h (step EUF); 6 - after ultrafiltration (step EPF); and 7 - concentrated extract after evaporation (step EBATCH). In each of these steps, four samples (repetitions) of 250 g each were collected and determined regarding vitamin C content, as triplicates.

For calibration group, samples consisted of four repetitions taken from each processing step plus another set of samples prepared after diluting the concentrated extract from step 7 (EBATCH) in water in the following proportions 4:0; 3:1; 2:2; 1:3; 0.5:3.5 and 0.25:3.75 ($v.v^{-1}$) that were equivalent to the concentration of 10,000 (C0); 7,500 (C1); 5,000 (C2); 2,500 (C3); 1,250 (C4) and 500 (C5) $mg.100\ g^{-1}$ of vitamin C, respectively. For the external validation group, five samples were taken from three stages, initial, intermediate and final of processing line.

### Total vitamin C determination

As an analytical reference method for constructing the calibration curve of total vitamin C content, the methodology of Chen and Wang (2000) was adapted. Thus, 2 mL of extract from each processing step were centrifuged at 15,000 x*g* for 15 min at 4 °C, afterwards, 0.1 mg of the supernatant was homogenized with 25 mL of 5% trichloroacetic acid (TCA), using a Vortex shaker (AP-56, Phoenix®, Brazil) for 20 s.

Thereafter, 50 µL aliquots were added to 25 µL of the 100 mM monobasic phosphate ($KH_2PO_4$) buffer, pH 7.0, and to 175 µL of a reaction solution containing 5 mL of 10% TCA, 5 mL of 44% phosphoric acid, 5 mL of 2,2'-bipyridyl (at 4% diluted in 70% ethanol) and 2.5 mL of 3% iron trichloride ($FeCl_3$) solution. A standard curve was prepared with aliquots of ascorbic acid (AsA, Sigma-Aldrich®) and 5% TCA in concentrations ranging from 0 to 50 µL, added to 25 µL of 100 mM $KH_2PO_4$ buffer, pH 7.0 and 175 µL of the above-mentioned reaction solution. The mixture was shaken for 5 s and incubated at 37 ºC for 60 min and then, analyzed by monitoring absorbance at 525 nm in a spectrophotometer (Synergy HT, Bio-Tek® Instruments, Inc., VT-USA) and results expressed in mg.100 $g^{-1}$.

**NIR scanning of samples**

Absorption data in the Near Infrared (NIR) region were obtained in 2 mL aliquots of the vitamin C extract using liquid sampler (Pkg 100) on a PerkinElmer® Frontier FT-IR/NIR Spectrum 100 N spectrometer (MA, USA), with temperature set at 18 (±2) ºC, with a configuration of 32 screenings in the range of 800 to 2,500 nm and a resolution of 4 $cm^{-1}$.

**Data analysis**

Spectral pretreatment and model construction were carried out using The Unscrambler X® version 10.4.1 (Camo Analytics, Norway) chemometric package. Pretreatments with Multiplicative Scatter Correction (MSC), Standard Normal Variation (SNV) and first derivative by Savitzky-Golay algorithm were evaluated. Multivariate regression models were developed using Partial Least Squares (PLS) regression and full cross-internal validation. Each regression model constructed was validated by samples external to the calibration group.

As recommended, ASTM E1655-05 standard practices (AMERICAN SOCIETY FOR TESTING AND MATERIALS, 2012) were used to stablish the procedures for quantification of chemical species using infrared spectroscopy. Performance of such models was evaluated using as parameters the coefficient of determination ($R^2$), the root mean square error of cross validation (RMSECV) and root mean square error of prediction (RMSEP).

## RESULT AND DISCUSSION

**Determination of vitamin C by the reference method**

Vitamin C content was determined by the reference analytical method during the stages of its extraction from immature acerola (Table 1). The mean vitamin C content of initial EAE step was 1708.23 mg.100 $g^{-1}$, while at the last EBATCH step, the mean content of 9969.35 mg.100 $g^{-1}$, approximately eight times more concentrated than previous steps. EBATCH is the final production step, thus the concentrated vitamin C extract is subsequently spray dried and encapsulated. Chewable tablets have very high vitamin C concentration as a strategy to avoid losses due to product degradation, in addition to providing practicality to commercialization (ABE-MATSUMOTO; SAMPAIO; BASTOS, 2018).

**NIR spectra**

Spectra of the samples collected during preparation of vitamin C extract of acerola ranged in absorbance from 800 to 2500 nm (Figure 1). In all extraction steps, prominent spectral bands were observed in regions with 1164; 1340; 1448; 1779; 1930; 2289 and 2492 nm of absorbance. NIR spectra were pre-processed with SNV and MSC to delete slope variation and to correct for scattering effects, once samples from EAE, EPE, EAH and DEC steps presented particles in suspension. Vitamin

**Table 1 -** Vitamin C determination during industrial extraction from immature acerola

| Steps* | Concentration range | Vitamin C (mg 100 $g^{-1}$) | | |
|---|---|---|---|---|
| | | Mean value | Standard error | Coefficient of Variation (CV) |
| EAE | 1678.11 – 1746.32 | 1726.80 | 32.55 | 1.9 |
| EPE | 1222.48 – 1306.36 | 1275.75 | 36.92 | 2.9 |
| EAH | 1774.75 – 1822.72 | 1808.54 | 22.80 | 1.3 |
| DEC | 1226.28 – 1368.48 | 1283.62 | 61.45 | 4.8 |
| EUF | 1440.31 – 1567.87 | 1512.86 | 60.43 | 4.0 |
| EPF | 1188.39 – 1444.57 | 1318.83 | 128.00 | 9.7 |
| EBATCH | 9809.08 – 9959.74 | 9905.23 | 67.32 | 0.7 |

*Extraction steps: EAE - before PME treatment; EPE - after PME treatment; EAH - after $CaOH_2$ addition; DEC - after decanting; EUF - during ultrafiltration; EPF - after ultrafiltration; and EBATCH - concentrated extract

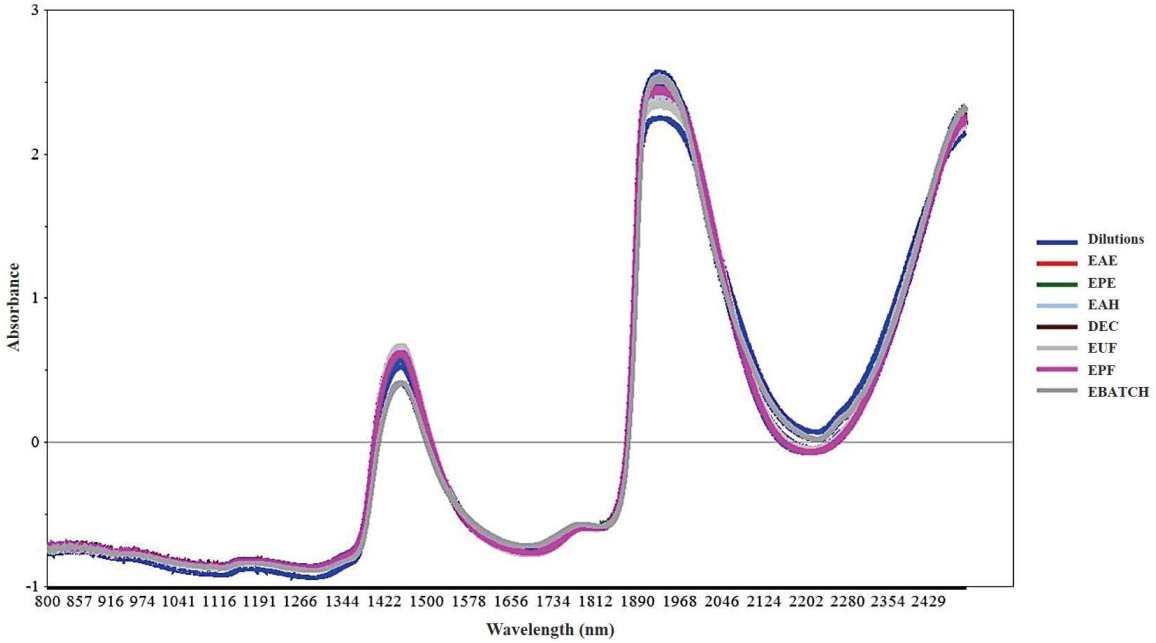
C functional groups were identified in the absorption bands of 1145-1202 nm and 1727-2020 nm, in addition to a peak at 1148 nm. Moraes *et al.* (2019), analyzed whole acerola with NIR and reported identification of vitamin C absorption bands between 1170-1180 nm, which corresponded to the second methylene group of overtones and a peak at 1410 nm that represented bands of combined methylene group of ascorbic acid, while acid absorption occurred between 1890 and 1950 nm.

Water absorption bands can be seen at 1400; 1775; 1900; 2275 and 2400 nm, although varied in absorbance values among samples from different extraction steps (Figure 1). The different absorbance values are related to changes in vitamin C content, as the final-step sample EBATCH presented the highest concentration. This result corroborates for the non-removal of water bands, once they did not interfere or mask vitamin C. In foods such as milk, potatoes, meat, tofu, wheat flour and rice, water bands were also identified in wavelength regions of 1300; 1700; 1900; 2200 and 2500 nm (NING-PFAUE, 2003).

In Figure 1, those bands shown at 1164; 1340; 1448; 1779; 1930; 2289 and 2492 nm and their vibrations in corresponding overtones may be seen in Table 2. The region from 950 to 1400 nm represents the first and second vibrations of water-related O-H stretching overtones (ARENDSE *et al*., 2018; MAGWAZA *et al*., 2012). The absorption band at 1930 nm corresponds to the second harmonic stretching vibration associated with C=O, in addition to combination of stretching and flexion vibration associated with O-H. Whereas, the region between 1200 and 1942 nm corresponds to first and second harmonic stretching vibrations associated with C-H, regardless of the presence of organic sugars. However, this region also includes the third harmonic vibration related to deformation of OH, CH and $CH_2$ bonds, found only in solutions containing sugars and organic acids (ARENDSE *et al*., 2018; GOLIC; WALSH; LAWSON, 2003). Sugars and organic acids were identified in wavelengths regions from 1100 to 1600 and from 1700 to 2300 nm (GÓMEZ; HE; PEREIRA, 2006; LOUW; THERON, 2010).

Therefore, Table 2 shows that vitamin C molecule has three functional groups: enol (–CH = CH – OH), ester (R'-COOR ") and alcohol (R – OH), whose overtones were observed in the spectrum (Figure 1). However and despite the fact that wavelength determination and band assignment have been carried out according to literature (ARENDSE *et al*., 2017, 2018; MAGWAZA *et al*., 2012), both physical and chemical nature of each sample should be taken into consideration as they may contribute to band displacement. Arendse *et al*. (2017), assigned wavelength regions from 1064 to 1333 nm and from 1640 to 2100 nm as corresponding to vitamin C of pomegranate processed in arils, while for whole pomegranate, the designated regions for vitamin C were from 1333 to 1640 nm and

**Figure 1 -** Spectral data from samples collected during vitamin C industrial extraction from immature acerola. Extraction steps: EAE - before PME treatment; EPE - after PME treatment; EAH - after $CaOH_2$ addition; DEC - after decanting; EUF - during ultrafiltration; EPF - after ultrafiltration; and EBATCH - concentrated extract



from 1836 to 2175 nm (ARENDSE *et al*., 2018). Garcia, Silva and Seixas (2013) reported bands at 1000; 1514 and 2082 nm for vitamin C in acerola processed pulp.

**Partial least squares regression models (PLS)**

Initially, an exploratory study was carried out to identify similarity among industrial extraction steps, thus principal component analysis (PCA) was employed using spectral data with pre-treatment of standard normal variate (SNV) transformation to eliminate multiplicative interferences of both scattering effects and particle size (Figure 2).

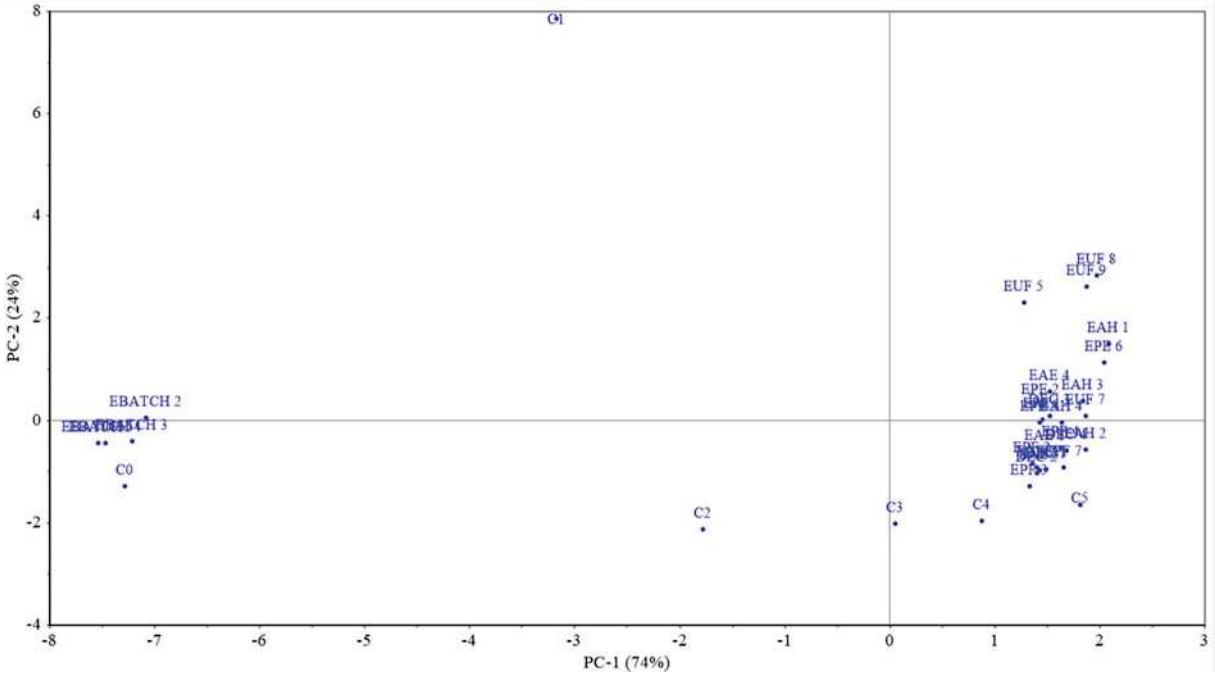
PCA confirmed the similarity in vitamin C content between samples from EAE, EAH, DEC and EPF steps, located to the right of the positive axis of PC1. Samples from EBATCH final step are located on the negative axis of PC1, distant from samples from other steps and between diluted samples originated from EBATCH with different concentration levels.

This result indicates how difficult it is to identify samples from specific extraction steps, with the exception of EBATCH. Thereby, models for predicting vitamin C quantification in three different steps have been proposed,

**Table 2** - NIR absorption bands of samples

| Wavelength (nm) | Vibrational modes |
|---|---|
| 1164 | C-H second harmonic stretching |
| 1340 | C-H stretching and flexion combined |
| 1448 | O-H first harmonic stretching |
| 1779 | C-H first harmonic tone stretching |
| 1930 | O-H stretching and flexion combined and C=O second harmonic stretching |
| 2289 | C-H stretching and flexion combined |
| 2492 | C-H, C-C combined and C-O-C stretching |

**Figure 2 –** Principal Component Analysis score plot of spectral data pretreated with standard normal variate transformation. Samples: EAE - before PME treatment; EPE - after PME treatment; EAH - after $CaOH_2$ addition; DEC - after decanting; EUF - during ultrafiltration; EPF - after ultrafiltration; and EBATCH - concentrated extract and diluted samples (C0 to C5)



the initial stage (EAE), after decantation (DEC) and the final concentrated extract (EBATCH). A model with only three of the extraction steps allows the identification of significant variation in vitamin C content during the process, thus allows possible corrections during its extraction and concentration procedures, and moreover, this may also be used for quality control of the final extract (EBATCH). The three selected steps have statistically different vitamin C concentrations, with 95% confidence, verified by the paired Student's t test with six freedom degree.

After steps selection, multivariate calibration models by partial least squares (PLS) algorithm were developed with different pre-treatments in the spectral data (Table 3). Together with PLS multivariate regression analysis, data pre-treatment methods were evaluated and among these, the best methods were SNV, MSC and first derivative of Savitzky-Golay using a window of 10 points. The performance model selected, within the employed pretreatment methods, was SNV for it had the highest $R^2$ value of 0.99 for calibration, and the lowest RMSECV value of 451.79 mg. 100 $g^{-1}$. Models were constituted of 34 samples (28 samples from extraction steps and 6 samples diluted from the EBATCH step). During modeling, 5 anomalous samples (outliers) were identified due to the high residue and influence on the model, therefore resulting in 29 samples for model calibration.

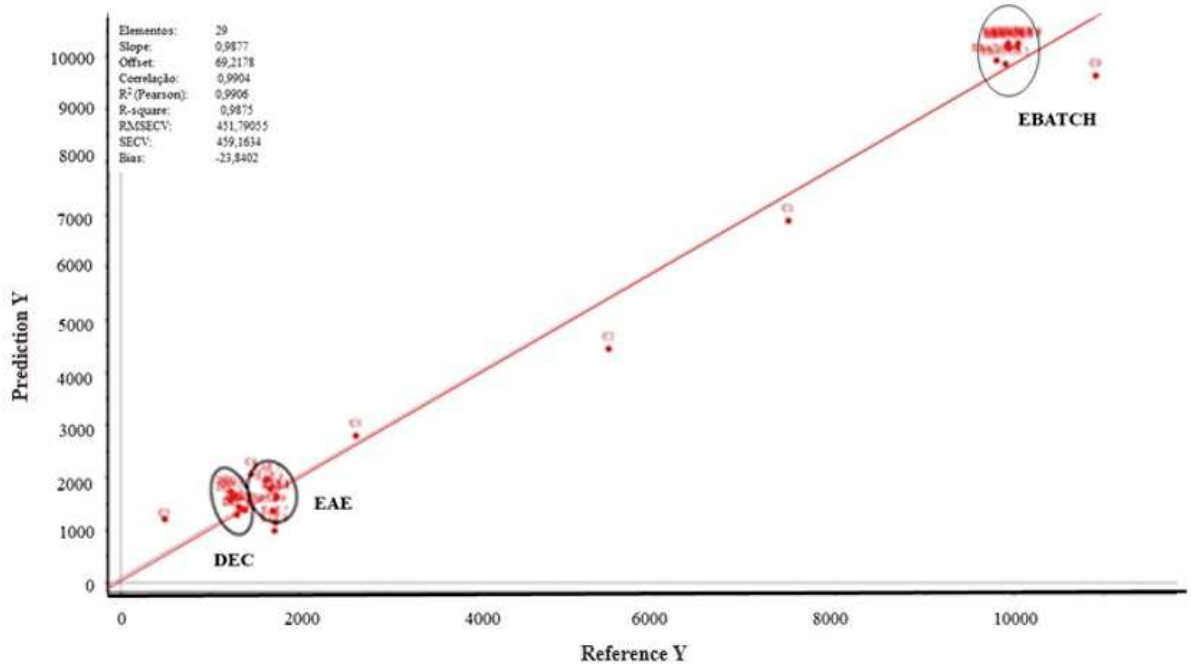
The best performance was found for PLS regression model and SNV pre-treatment showed (Figure 3) with a good definition for samples from three steps of vitamin C extraction from acerola. Samples

**Table 3 -** Performance of PLS regression model for vitamin C determination using NIR spectra

| Pretreatment | Calibration (n = 29) | | | $Rv^2$ | Cross Validation (n = 29) | | | |
|---|---|---|---|---|---|---|---|---|
| | LV | $Rc^2$ | RMSEC | | RMSECV | Bias | Slope | Corr |
| SNV | 2 | 0.99 | 406.37 | 0.99 | 451.79 | -23.84 | 0.98 | 0.99 |
| MSC | 2 | 0.99 | 406.74 | 0.99 | 452.37 | -23.90 | 0.98 | 0.99 |
| 1stDERIV | 3 | 0.99 | 376.94 | 0.98 | 498.10 | -10.54 | 0.97 | 0.99 |

*LV, latent variables; $R^2$, determination coefficient; RMSEC, root mean square error of calibration; RMSECV, root mean square error of calibration of internal validation; Bias, systematic variation error; Slope, steepness of line; Corr, correlation

**Figure 3** - Prediction x Reference of internal validation (n=29) with samples from EAE, DEC, EBATCH extraction vitamin C steps and samples diluted from EBATCH (C0 to C5)



from the initial processing step EAE are located between concentrations C4 and C3 of the diluted solutions (1250 to 2500 mg. $100g^{-1}$), samples from the intermediate processing step DEC are located between concentrations C5 and C4 of the diluted solutions (500 to 1250 mg $100g^{-1}$) while samples from final step EBATCH are located in proximity of C0 concentration (10000 mg. $100g^{-1}$).

To verify the potential of the model in predicting future samples, we simulated the use of NIR spectroscopy in the production line of vitamin C extract from acerola. With this purpose, a new PLS model was built with 24 samples (in accordance with ASTM E1655- 05, which determines a minimum of 24 samples for multivariate calibration) and prediction of 5 samples taken from three extraction steps (Table 4). Parameters of the calibration model and prediction of 5 samples showed RMSEP of 166.27 mg.$100g^{-1}$, which is smaller than obtained in the cross validation with excellent correlation between predicted values and those determined by the reference method.

Table 5 shows a comparison between values by NIR model and those determined by reference method, and the average relative error observed was lower than 5%, indicating a promising result for NIR spectroscopy use in quality control during vitamin C industrial extraction from immature acerolas.

To develop models, calibration, validation and prediction data were based on data from spectrophotometric reference analysis (Table 1) that were in a normal distribution. Reference data is the basis for a robust model and its interpretation depends on data accuracy and variation to obtain a satisfactory set and calibration (ARENDSE *et al*., 2018; LU *et al*., 2006), thus large variation within samples interferes directly with prediction model response (MAGWAZA *et al*., 2012).

The low CV values between 1.0 to 8.7 (Table 1) with the high RMSECV value of 451.79 mg.100 $g^{-1}$ obtained by PLS regression model (Table 3) indicate a low variability in vitamin C content among samples during its extraction from acerola. Thereby, the concentrated EBATCH sample was diluted in water to prepare a standard calibration curve, increasing variability and prediction in order to decrease RMSEC and RMSEP values for all models, so that $R^2$ and correlation values obtained were greater than 97%. The methods proved to be robust once the high RMSECV (451.79 mg 100 $g^{-1}$) and RMSEP (166.27 mg 100 $g^{-1}$) values are still small when compared to the average vitamin C content of the most concentrated EBATCH sample (9,969.35 mg. 100 $g^{-1}$), in addition to presenting $R^2$ values greater than 0.98 for calibration and 0.99 for validation.

Malegori *et al*. (2017), evaluated the vitamin C content in processed pulp from physiologically mature *acerola cv. Junko* using a Perkin-Elmer FT-NIR spectrometer and the reference analysis resulted in mean content of 2844 mg. 100 $g^{-1}$ with minimum and maximum values of 1576 and 3653 mg.100 $g^{-1}$, respectively. Results obtained by PLS model showed RMSEC of 145 mg. 100 $g^{-1}$ and RMSEP of 360 mg. 100 $g^{-1}$ and the authors supported their model validity based on RMSEC with $R^2$ of 0.91 and RMSEP with $R^2$ of 0.71. Garcia, Silva and Seixas (2013), to avoid a low CV value, opted for enriching acerola pulp samples with ascorbic acid ensuring a significant

**Table 4 -** Summary of PLS regression model for the prediction of vitamin C content using NIR spectra

| | Calibration (n = 24) | | | $Rv^2$ | Prediction (n = 5) | | | |
|---|---|---|---|---|---|---|---|---|
| | LV | $Rc^2$ | RMSEC | | RMSEP | Bias | Slope | Corr |
| PLS model | 2 | 0,99 | 405,58 | 0,99 | 166,27 | 34,67 | 1,04 | 0,99 |

*LV, latent variables; $R^2$, determination coefficient; RMSEC, root mean square error of calibration; RMSECV, root mean square error of calibration of internal validation; Bias, systematic variation error; Slope, steepness of line; Corr, correlation

**Table 5 –** Comparison of vitamin C content in predicted samples with relative error between -7.70% and 6.88%

| Samples | NIR model mg 100 $g^{-1}$ | Reference method mg 100 $g^{-1}$ | Relative Error % |
|---|---|---|---|
| EAE 1 | 1608.10 | 1742.18 | -7.70 |
| EAE 2 | 1636.18 | 1746.32 | -6.31 |
| DEC 1 | 1415.96 | 1324.87 | 6.88 |
| DEC 2 | 1379.17 | 1368.48 | 0.78 |
| EBATCH 1 | 10257.84 | 9942.06 | 3.18 |

correlation between the reference and prediction methods. An increase in variability due to ascorbic acid enrichment in mandarin juice samples resulting in a higher CV was also reported by Liu, Chen and Ouyang (2008).

Moraes *et al*. (2019), analyzed whole acerola at different maturation stages with FT-IR/NIR using pretreatments MSC, first derivative and PLS model able to predict vitamin C concentration in those fruits. The authors reported RMSECV of 402.40 mg. 100 $g^{-1}$ and RMSEP of 355.20 mg. 100 $g^{-1}$, which did not perform well for the prediction model, however, data was submitted to GA-PLS model (genetic algorithm combined with partial least squares) and after data processing with the second derivative, a better model was obtained with RMSECV of 22.90 mg. 100 $g^{-1}$and RMSEP of 46.30 mg. 100 $g^{-1}$. Finally, authors explained the genetic variation observed in data justified the use of GA-PLS, which predicts mutations and descendants within the population as may be the case with whole fruits.

The prediction error (RMSEP 166.27 mg 100 $g^{-1}$) obtained in this work, although higher than the reference method error, corroborates with the use of NIR spectroscopy as a promising tool for monitoring vitamin C content during extraction process from acerola, once the possibility of measuring its content in intermediate (DEC) and final (EBATCH) steps of the production line, enables necessary adjustments to be made.

## CONCLUSIONS

1. During industrial extracting of vitamin C from acerola, the processing steps have statistical similarities with 95% confidence regarding vitamin C content;
2. However, three steps stood out in the process, the initial EAE, the intermediate DEC and the final concentrated EBATCH, and the use of NIR spectroscopy associated with PLS model algorithm showed good results for prediction of vitamin C content in these steps with relative errors of less than 5%;
3. These results corroborate with development of prediction models employing NIR spectroscopy to quantify vitamin C extracted from acerola, which may be used by the industry for a fast and effective quality control system during extraction with shorter sample-preparation time, free of waste and environmentally friendly.

## AKNOWLEDGEMENTS

We thank Nutrilite™ da Amway do Brasil for the samples used in our study, Nuteral Indústria de Formulações Nutricionais Ltda for allowing us to use their NIR spectrometer, Institutos Nacionais de Ciências e Tecnologia Analíticas Avançadas (INCTAA, CNPq N° 465.768/2014-8) and Frutos Tropicais (INCT-FT) and CNPq, Brasil for the scholarships and grants.

## REFERENCES

ABE-MATSOMOTO, L. T.; SAMPAIO, G. R.; BASTOS, D. H. M. Do the labels of vitamin A, C, and E supplements reflect actual vitamin content in commercial supplements? **Journal of Food Composition and Analysis**, v. 72, p. 141-149, 2018.

ALAMAR, P. D. *et al.* Quality evaluation of frozen guava and yellow passion fruit pulps by NIR spectroscopy and chemometrics. **Food Research International**, v. 85, p. 209-214, 2016.

AMERICAN SOCIETY FOR TESTING AND MATERIALS. **E1655-05**: standard practices for infrared multivariate quantitative analysis. West Conshohocken: ASTM, 2012.

AMODIO, M. L. *et al.* Potential of NIR spectroscopy for predicting internal quality and discriminating among strawberry fruits from different production systems. **Postharvest Biology and Technology**, v. 125, n. 125, p. 112-121, 2017.

ARENDSE, E. *et al*. Development of calibration models for the evaluation of pomegranate aril quality by Fourier-transform near infrared spectroscopy. **Biosystems Engineering**, v. 159, p. 22-32, 2017.

ARENDSE, E. *et al.* Fourier transform near infrared diffuse reflectance spectroscopy and two spectral acquisition modes for evaluation of external and internal quality of intact pomegranate fruit. **Postharvest Biology and Technology**, v. 138, p. 91-98, 2018.

BELWAL, T. *et al.* Phytopharmacology of Acerola (*Malpighia* spp.) and its potential as functional food. **Trends in Food Science and Technology**, v. 74, p. 99-106, 2018.

CALGARO, M.; BRAGA, M. B. **A cultura da acerola**. 3. ed. rev. ampl. Brasília, DF: Embrapa, 2012. 144 p.

CARAMÊS, E. T. S. *et al*. Quality control of cashew apple and guava nectar by near infrared spectroscopy. **Journal of Food Composition and Analysis**, n. 56, p. 41-46, 2017.

CASALE, M. *et al*. NIR, spectroscopy-based efficient approach to detect fraudulent additions within mixtures of dried porcini mushrooms. **Talanta**, v. 160. p. 729-734, 2016.

CHEN, J.; WANG, X. Experimental instruction of plant physiology. **South China University of Technology Press**, p. 124, 2000.

GARCIA, V. A. S.; SILVA, M. R.; SEIXAS, F. A. V. Rapid analysis of vitamin C content in acerola extract by FT–NIR spectroscopy. **Revista Tecnológica**, v. 22, p. 13-21, 2013.

GOLIC, M.; WALSH, K. B.; LAWSON, P. Short-wavelength near infrared spectra of sucrose, glucose, and fructose with respect to sugar concentration and temperature. **Applied Spectroscopy**, v. 57, p. 139-145, 2003.

GÓMEZ, A. H.; HE, Y.; PEREIRA, A. G. Non-destructive measurement of acidity, soluble solids and firmness of Satsuma mandarin using vis–NIR spectroscopy techniques. **Journal Food Engineering**, v. 77, p. 313-319, 2006.

HU, J. *et al.* Rapid evaluation of the quality of chestnuts using near-infrared reflectance spectroscopy. **Food Chemistry**, n. 231, p. 141-147, 2017.

LIU, Y.; CHEN, X.; OUYANG, A. Nondestructive determination of pear internal quality indices by visible and near-infrared spectrometry. **Food Science and Technology**, v. 41, p. 1720-1725, 2008.

LOUW, E. D.; THERON, K. I. Robust prediction models for quality parameters in Japanese plums (*Prunus salicina* L.) using NIR spectroscopy. **Postharvest Biology and Technology**, v. 58, p. 176-184, 2010.

LU, H. *et al.* Application of Fourier transform near infrared spectrometer in rapid estimation of soluble solids content of intact citrus fruits. **Journal of Zhejiang University Science**, v. 7, p. 794-799, 2006.

MAGWAZA, L. S. *et al.* NIR spectroscopy applications for internal and external quality analysis of citrus fruit: a review. **Food and Bioprocess Technology**, v. 5, p. 425-444, 2012.

MALEGORI, C. *et al*. Comparing the analytical performances of Micro-NIR and FT-NIR spectrometers in the evaluation of acerola fruit quality, using PLS and SVM regression algorithms. **Talanta**, v. 165, p. 112-116, 2017.

MORAES, F. P. *et al*. Estimation of ascorbic acid in intact acerola *(Malpighia emarginata* DC) Fruit by NIRS and Chemometric Analysis. **Horticulturae**, v. 5, n. 1, p. 12, 2019.

NING-PFAUE, H. B. Analysis of water in food by near infrared spectroscopy. **Food Chemistry**, v. 82, p. 107-115, 2003.

SNYDER, A. B. *et al.* Rapid authentication of concord juice concentration in a grape juice blend using Fourier-Transform infrared spectroscopy and chemometric analysis. **Food Chemistry**, v. 147, n. 2014, p. 295-301, 2014.

TIERNO, R. *et al*. Phytochemicals determination and classification in purple and red-fleshed potato tubers by analytical methods and near infrared spectroscopy. **Journal of the Science of Food and Agriculture**, v. 96, n. 6, p. 1888-1899, 2016.

TOLEDO-MARTÍN, E. M. *et al.* Application of visible/near-infrared reflectance spectroscopy for predicting internal and external quality in pepper. **Journal of the Science of Food and Agriculture**, v. 96, n. 9, p. 3114-3125, 2016.

VENDRAMINI, A. L.; TRUGO, L. C. Chemical composition of acerola fruit (*Malpighia punicifolia L.*) at three stages of maturity. **Food Chemistry**, v. 71, n. 2, p. 195-198, 2000.