Postharvest Biology and Technology 141 (2018) 86–97

Contents lists available at ScienceDirect

Postharvest Biology and Technology

journal homepage: www.elsevier.com/locate/postharvbio

Validation of short wave near infrared calibration models for the quality and ripening of 'Newhall' orange on tree across years and orchards

Ana M. Cavaco[a,*], Rosa Pires[a], M. Dulce Antunes[a,b], Thomas Panagopoulos[c], António Brázio[a], Andreia M. Afonso[a], Leonardo Silva[a], Márcia Rosendo Lucas[d], Bernardo Cadeiras[d], Sandra P. Cruz[a], Rui Guerra[a]

[a] CEOT, Universidade do Algarve, Campus de Gambelas, 8005-189 Faro, Portugal
[b] MeditBio, Universidade do Algarve, Campus de Gambelas, 8005-189 Faro, Portugal
[c] CIEO, 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:
Citrus
SW-NIRS
Fruit quality attributes
On-tree
Model robustness
Precision agriculture

ABSTRACT

The aim of this research was to test the viability of short wave near infrared spectroscopy (SW-NIRS) for the monitoring of fruit quality and ripening evolution in Algarve Citrus orchards (*Citrus sinensis* L. Osbeck 'Newhall'). Specifically, we have investigated the robustness of SW-NIRS calibration models in real conditions, that is: *i)* measurements were performed on tree, at a single location in the fruit equator, in the sunlight and with no temperature equilibration; and *ii)* with validation through independent data obtained in different years and/or orchards. Calibration models for soluble solids content (SSC), juice pH, titratable acidity (TA), firmness and maturation index (MI = SSC/TA) were built from the spectral data obtained in two orchards with different edaphoclimatic conditions, and in two consecutive years, corresponding to four independent datasets. We propose a method to assess model robustness through the comparison of internal validation (IV: calibration and validation data sets homogeneously sampled from the whole data set) and external validation (EV: calibration and validation data sets corresponding to different orchards and/or years). The method is based on the statistics of the results obtained by either IV and EV when applied to all the possible combinations of the four datasets. The results show that IV overestimates the models' performance relatively to the realistic exercise of EV. Globally, SSC and juice pH were the best performing models, with fair performances in IV and poor performances in EV (example for SSC: (IV/EV): rmsep = 1.00/1.15%, SDR = 1.40/1.13, R^2 = 0.49/0.34). Firmness yielded the worse models, while TA and MI yielded intermediate performances. However, the plots derived from the comparison method suggest a convergence of IV and EV performances for larger numbers of samples, and thus the potential for future continuous model improvement.

1. Introduction

Citrus fruit are one of the most valuable crop in international trade (Ladaniya, 2008). Algarve Citrus ("Citrinos do Algarve") are protected geographical indication (PGI) fruit with a major economic impact both in the national market and in the exports shares of Portugal. As non-climacteric fruit, citrus is only harvested at their optimal eating ripening stage. Algarve Citrus must further observe the requirements stipulated by the respective PGI norm (Uniprofrutal, 2003). Additionally, to the mandatory sanitary conditions and external aspect of the fruit required for handling, transport, storage and a proper shelf-life, the decision on Algarve Citrus optimal harvest date (OHD) is based on: i) soluble solids content (SSC); ii) fruit juice percentage; and iii) maturity index (MI) [SSC/ titratable acidity (TA)], which differ among the various species/varieties/hybrids comprised by the PGI.

The internal quality attributes (IQA) change along fruit ripening and must be determined almost weekly close to harvest. Their value depends on the cultivar, year, orchard and tree site, and fruit location on the tree. The usual approach to establish OHD is to collect fruit sets from each orchard by the beginning of the harvest season and use them to determine the mandatory IQA by standard destructive and time-consuming methods. The number of measured fruit is not statistical representative, resulting in a fallible evaluation of IQA and orchards maturity, deficient crop management and lack of transparency in the citrus supply chain. This is a main concern for traceability, which has become an important issue (Beulens et al., 2005).

* Corresponding author.
E-mail address: acavaco@ualg.pt (A.M. Cavaco).

https://doi.org/10.1016/j.postharvbio.2018.03.013
Received 30 November 2017; Received in revised form 10 March 2018; Accepted 16 March 2018
Available online 26 March 2018

Overall, the lack of the technology for a fast, extensive and reliable determination and record of the fruit IQA at the outmost critical step of the ripening process, will have a toll on every step along the postharvest quality of the Algarve Citrus supply chain. Such technology should allow to forecast the OHD and thus to optimize the management of the orchards, a complex task running throughout the whole year, due to the many species/varieties/hybrids comprised in Algarve Citrus.

Visible near-infrared spectroscopy (Vis-NIRS) is non-destructive, fast, objective, flexible and versatile. It has proved to predict effectively many ripening and quality attributes of a wide number of fruit including citrus (Guthrie et al., 2005a,b; Nicolaï et al., 2007; Cayuela, 2008; Magwaza, et al., 2012a,b; Magwaza et al., 2013, Magwaza et al., 2014a,b,c; Sánchez et al., 2013a,b; Liu et al., 2015; Bizzani et al., 2017; Ncama et al., 2017; Torres et al., 2017, Arendse et al., 2018).

Most calibration models published for citrus, as well as other fruit, are still based on samples collected and assessed under controlled conditions in the laboratory after fruit temperature equilibration. Despite the availability of portable spectrometers has further opened the possibility of using this technology on the field, only a few studies have focused on the use of this technology to assess the quality and ripening attributes of oranges and mandarins on-tree, perhaps due to the complexities involved (Zude et al., 2008; Sánchez et al., 2013a, b; Torres et al., 2017).

An emerging research trend in the application of Vis-NIRS is to improve model robustness that must be accounted for if one wants this technology to be incorporated on the daily routines of producers and upgrade the current procedures to that of a precision agriculture approach. Vis-NIRS predictions are based on chemometrics models that must stand validation against unknown samples. Laboratory models with homogeneous fruit sets are abundant in the literature, but stringent multi-year, multi-cultivar and multi-orchard validations are much more difficult to perform and to find in the current literature (Magwaza et al., 2012a). Nevertheless, there are some clear examples of this approach, such as previously reported for mandarin (Guthrie et al., 2005b) and oranges (Zude et al., 2008; Cayuela and Weiland, 2010; Magwaza et al., 2013).

The most common procedure when constructing and validating calibration models for the various fruit quality attributes is to separate a fraction of the available samples as validation set (where one performs cross validation to build the model), and the remaining as calibration set. Furthermore, the validation samples are typically chosen as the best possible representation of the whole set. This is indeed the correct formal approach to demonstrate that a given model describes a given population. This has been applied even when the models comprehend several species and/or cultivars, orchard locations and harvest years, which are mixed in the calibration and validation sets (e.g. Torres et al., 2017). Yet, this is a static or closed approach - we will call it "internal validation" (abbreviated by IV) – because it is not challenged by an external dataset. Internal validation does not insure the success of a continuous monitoring application, which is a dynamic and open process. In this type of approach, the validation samples come from the following year or from a new orchard located in a different area, not from a closed dataset.

A truly stringent external validation (EV) is thus required to have a realistic idea of the models' performance in orchard monitoring. EV means validation through a dataset with a different origin (spatial or temporal) relatively to the datasets used in calibration. The first one was performed by Miyamoto and Kitano (1995), who have applied a SSC calibration model deployed for 'Satsuma' mandarins in one season to the subsequent two seasons. He succeeded to attain similar statistics to those obtained in calibrations developed within a given season (rmsep $<$ 0.6% SSC and bias $\leq$ 0.4% SSC). However, the measurements were not performed on-tree. Otherwise, Ou et al. (1997) when performed inter-orchard validation of a model for SSC in 'Ponkan' mandarin for one region and applying it to other two regions, found significant degradation of the model performance (rmsep increased 70–90 % from the base value of 0.68% SSC). As expected, the combination of all data lead to a better model, minimizing the performance degradation to 35% (0.92% SSC). Guthrie et al. (2005a, b) used several data sets of different locations, harvest years and orchards to calibrate and validate the models for SSC and DM of 'Imperial' mandarins. A significant decrease of the accuracy of the calibration models was observed for both parameters, this depending on the calibration model fruit set and the validation population used. For instance in the multi-orchard calibration model comprising fourteen fruit independent populations, the correlation varied between $Rc^2 = 0.84$ and $Rv^2 = 0.59$. They concluded that a lack of robustness was obvious in terms of the ability of the models to predict attribute levels, namely SSC and DM in new and independent populations. Furthermore, Zude et al. (2008) used data sets of five orange varieties collected on-tree, from a commercial orchard located in California, in two different harvest seasons to calibrate and validate a multi-year/multi-variety prediction model for SSC. Validation of the calibration model for SSC obtained in the season 2004/2005 with data obtained in 2005/2006, showed a clear degradation of the model performance, with a significant increase of rmsep from ~0.86% SSC to 4.56% SSC). As previously shown by Guthrie et al. (2005b), this negative effect was compensated through a re-calibration of the first season model, using only a partial data set of 2005/2006, which decreased the error to the initial value.

Although the recalibration procedure is feasible if one may assume that the sampled fruits used to recalibrate the model constitute a faithful representation of the new validation population, this becomes quite difficult to apply if one aims to monitor the fruit evolution through time. For example, the fruit sampled in the first weeks cannot represent those to be measured in the last weeks of the harvest season. Thus, it is not possible to recalibrate the models for an entire season based on recalibration samples acquired at its beginning. Overall, the recalibration should be performed regularly along all the new season, and this would be impractical.

The present investigation is included in a long-term project that aims to assess the viability of short wave near infrared spectroscopy (SW-NIRS) for the monitoring of fruit quality and ripening evolution in several Algarve Citrus orchards by assessing fruit on-tree under field and real-life conditions along the harvest season and to provide a smart tool to determine the OHD of each variety. As such, the objective of this study was to test robustness and reproducibility of the calibration model's predictions across years and geographical sites, based on the evaluation of the application of SW-NIRS in totally real conditions. The fundamental difference relatively to previous works is that we use a comparison of IV and EV as a tool to investigate model robustness. Also, the conditions were the most realistic possible: spectra were taken *in situ*, from 'Newhall' oranges on-tree, without any temperature equilibration, fruit temperature changing from 11 °C to 27 °C (Table 2). Obviously, this is expected to result in the degradation of the model performance because the spectra are dependent on temperature (Peirs et al., 2003). Probing the fruit in the tree also brings reproducibility issues. However, we are assessing the viability of the SW-NIRS models' performance in real conditions and this is unavoidable.

2. Material and methods

2.1. Fruit

The oranges (*Citrus sinensis* L. Osbeck 'Newhall') used in these tests were picked randomly at the eye level (circa 1.60 m height) of the canopy of each of the 25 geo-referenced trees chosen across two commercial orchards of CACIAL, located in Quarteira (37°04′50.93″ N 8°04′05.66″ O; elevation 26 m) and in Paderne (37°11′07.73″ N 8°10′44.74″ O; elevation 65 m), South Portugal. Sampling was performed through time, starting from early ripening stage up to late harvest, in both orchards, in the following periods: October 2015-February 2016 (harvest season 1) and November 2016-February 2017