Computers and Electronics in Agriculture 155 (2018) 23–31

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Determination of soluble solid content in multi-origin 'Fuji' apples by using FT-NIR spectroscopy and an origin discriminant strategy

Xiaona Li, Jichao Huang, Yingjun Xiong, Jun Zhou, Xiangyu Tan, Baohua Zhang*

College of Engineering, Nanjing Agricultural University, Nanjing 210031, Jiangsu, PR China

ARTICLE INFO

Keywords:
Spectroscopy
Multi-origin model
SSC prediction
Geographical origin
Model search strategy

ABSTRACT

Apple is widely planted all over the world. Origin variability influences the internal quality of apples because soil characteristics, light effects, nutrition, weather conditions, as well as growing management vary from orchard to orchard. However, if taking spectral variations caused by the origin variability of apple samples from different orchards into account, the fruit quality parameters could not be measured or predicted with high accuracy by using the current models without updates. To improve the practicability and accuracy of the prediction models, a multi-origin regression model for the determination of soluble solids content in apples from three origins by using FT-NIR spectroscopy and a model search strategy was developed in this paper. In this model, based on the wavelengths selected by competitive adaptive reweighted sampling algorithm (CARS), partial least squares discriminant analysis (PLS-DA) was trained and applied to identify the geographical origins of the apple samples. The results indicate that the samples spectra were correctly matched to the corresponding classes and a 98.1% correct classification was achieved. Partial least squares regression (PLS) was used to establish three single-origin calibration models for the determination of soluble solids content (SSC) in apples from three different origins, and meanwhile, CARS algorithm was also applied to select the most effective wavelengths for calibration models. Then, the multi-origin CARS-PLS model for determination of SSC in apples from three origins was developed combined with origin discriminant and the proposed model search strategy. It was concluded that the multi-origin CARS-PLS model achieved more satisfying results than the single-origin CARS-PLS models for the determination of SSC, with R_p and *RMSEP* values for the apple samples from three geographical origins being 0.921, 0.759, 0.924 and 0.661, 0.673, 0.547 Â°Brix, respectively. The above results indicate that it is promising to build a multi-origin CARS-PLS model to predict SSC for apples based on an origin discriminant approach to reduce the effect of geographical origin.

1. Introduction

Apple is a popularly produced and consumed fruit which is highly favored by the consumer because of its rich nutrition and healthcare benefits (Filipic, 2011). Apple is also a good source of antioxidant components, such as ascorbic acid and polyphenolic compounds, which can exert protective effects against various degenerative diseases (Giovanelli et al., 2014; Fan et al., 2016). External appearance and internal quality of apples are critical attributes that directly influence the purchasing decision of consumers on apple fruit. SSC is a key parameter to evaluate the internal quality of apple, and also is a comprehensive index including soluble sugar, acid, cellulose, and other ingredients (Mendoza et al., 2011). Therefore, in an attempt to meet the growing consumer demands for high quality of apples, it is crucial to develop a rapid and reliable SSC detection method for apple. Most instrumental techniques to measure SSC used to be destructive, and involved a considerable amount of manual work. However, compared with traditional destructive testing methods, the use of near-infrared (NIR) spectroscopy to measure quality characteristics fruit during the last decade, especially SSC, because it is precise, reliable, rapid, non-destructive and inexpensive (Fan et al., 2009).

NIR spectroscopy is a spectroscopic approach that covers the region of the electromagnetic spectrum between 780 and 2500 nm (Lin and Ying, 2009). The recorded NIR spectroscopy contains both physical and chemical information of the radiated samples (Zhang et al., 2014). Combined with an appropriate predictive model which could be established by partial least square (PLS), principal component regression (PCR) and other various statistical methods (Zhang et al., 2017), NIR spectroscopy has been demonstrated to be successfully utilized for non-destructive SSC assessment of apples (Xie et al., 2008; Ngouajio et al.,

* Corresponding author.
E-mail address: bhzhang@njau.edu.cn (B. Zhang).

https://doi.org/10.1016/j.compag.2018.10.003
Received 26 July 2018; Received in revised form 30 September 2018; Accepted 3 October 2018

2007). For instance, Peirs et al. (2003) analyzed the influence of temperature fluctuations on the near-infrared spectrum detection of SSC in apples. Zou et al. (2007) investigated the feasibility of rapidly measuring intact apple SSC with a fiber optic Fourier transform near-infrared (FT-NIR) spectrometer. More applications of FT-NIR spectrometer and recent development have been briefly reviewed by Peirs et al. (2002) and Liu et al. (2005).

Apple is a widely cultivated fruit crop with a wide range of origins. Due to the large span of growing areas and variation in growing environments, there is a large difference in the external and internal quality of the apples produced in different origins (Tian et al., 2017). Bobelyn et al. (2010) investigated the effects of origin variability on the measured spectra and evaluated its effect on the performance of NIR calibration models to predict SSC. The origin variability influences the composition and quality of fruit because that soil nutrient, light effects, climate and growing management are variable from orchard to orchard (Zhang et al., 2015). Consequently, this biological variability will affect the prediction of the quality parameters and also the developed models.

Owing to the variation in quality of the fruit from different origins, the internal quality parameters of the apple samples from a specific origin only represent local fruit, and can't represent the fruit from other growing areas with different growing conditions. As origin variability is not taken into account, the higher prediction errors of the SSC assessment model will be obtained. As a consequence, the methods used in Bobelyn et al. (2010) needed to be improved or optimized to detect SSC of the fruit from different origins. However, there were few reports on eliminating the effect of origin variability on the accuracy of detection models for SSC. Fan et al. (2015) explored the influence of the different origins in the internal quality inspection by using spectral information and provided a solution by using a hybrid origin model. In their study, to eliminate the effect of origin variability, the hybrid origin model was developed by the full samples data from all different origins. It was found that the hybrid origin model showed good ability in reducing the origin effect on the robustness of prediction models for SSC of apple. However, the hybrid origin model needs to be updated when applied to the future samples from more different origins, its prediction accuracy will decrease as the origins increase. Therefore, to eliminate the effect of origin variability on NIR model without reducing the prediction accuracy, our manuscript provided a new routine for solving this problem by developing a multi-origin regression model.

2. Materials and methods

2.1. Samples feature

In this study, the samples were purchased from 'Fuji' apple commercial orchards. A total of 208 samples with no defects and damages were purchased from three different 'Fuji' apple commercial orchards, of which 76 apple samples were from Akesu, 72 from Qixia, and 60 from Yichuan. The elevations of Akesu, Qixia, and Yichuan are respectively 1108 m, 149 m, and 836 m. As elevation increases, climate changes and the air becomes colder and drier, having a corresponding impact on plant growth. Soil characteristics, light effects, nutrition, weather conditions, as well as management practices, vary from location to location.

All samples were individually washed, dried, numbered, and then sampling point marked around the equator. Before being examined by FT-NIR technique, apples were stored in the laboratory (Temperature, 20 °C; relative humidity, 60%) for 24 h to make the samples reaching room temperature and reduce the effect of samples temperature on the prediction accuracy (Tian et al., 2017). All measurements including spectral collection and quality analysis were carried out on the same day.

The samples were sorted into the calibration set and prediction set according to a proportion of 3:1 (Hu et al., 2015). Therefore, a subset of 156 samples (57, 54 and 45 for samples from Akesu, Qixia and Yichuan, respectively) were selected as calibration set for developing the calibration model, while the remaining 52 samples (19, 18 and 15 for samples from Akesu, Qixia, and Yichuan, respectively) were used as a prediction set to verify the prediction performance of the calibration model. To compare the performance of different calibration models, the samples in the calibration and prediction sets remained unchanged for all models.

2.2. FT-NIR spectra collection

The spectral data of apple samples were collected by an Antaris II FT-NIR spectrometer (Thermo Electron, USA), fitted with an integrating sphere, a high sensitivity InGaAs detector, and a tungsten lamp (20 W) to provide diffuse reflectance measurement. The bifurcated optical cable and the fruit holder were the main accessories in the experiment. In the head of the bifurcated cable, the source and detector fibers were situated randomly (Liu et al., 2005). Apples were placed steadily on the fruit holder, with the equidistant horizontal. The light was guided to the sample by source fibers, then reflected back from the sample and received by the detector fibers and delivered to the FT-NIR spectrometer. The integration sphere is used to recover the fraction of radiation that is not absorbed from the sample and reflected in the direction of the internal surface of the sphere, retrieving the important information that may be lost. The schematic diagram of the setup for FT-NIR measurement of SSC in apple is shown in Fig. 1. The FT-NIR spectrometer had a spectral range of 10,000 and 4000 cm^{-1} with a 1.928 cm^{-1} spectral

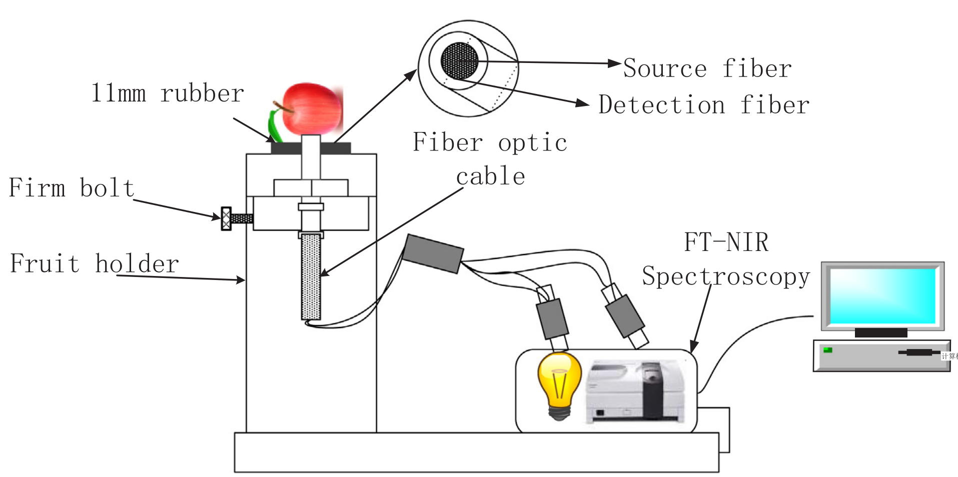

Fig. 1. Schematic diagram of the setup for FT-NIR measurement of SSC in apple.