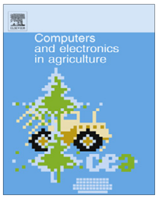
Computers and Electronics in Agriculture 140 (2017) 361–373

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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



Original papers

Chemical imaging for measuring the time series variations of tuber dry matter and starch concentration

Wen-Hao Su, Da-Wen Sun [*]

Food Refrigeration and Computerised Food Technology (FRCFT), School of Biosystems and Food Engineering, Agriculture & Food Science Centre, University College Dublin (UCD), National University of Ireland, Belfield, Dublin 4, Ireland

ARTICLE INFO

Article history:
Received 2 August 2016
Received in revised form 10 June 2017
Accepted 10 June 2017
Available online 22 June 2017

Keywords:
Spectral imaging
Multivariate regression
Wavelength selection
Visualization
Potato
Sweet potato

ABSTRACT

The potential of chemical imaging for rapid measurement of dry matter concentration (DMC) and starch concentration (SC) in both potato and sweet potato tubers was investigated. The time series images of tuber samples were acquired, then the resulting reflectance spectra (R_S) were corrected and transformed into absorbance spectra (A_S), and exponent spectra (E_S). Full wavelength regression models including multiple linear regression (MLR), partial least squares regression (PLSR) and locally weighted partial least squares regression (LWPLSR) were established based on spectral profiles with measured DMC and SC values. The best calibration model for measuring DMC and SC was LWPLSR based on E_S and R_S where the coefficients of determination in cross-validation (R^2_{CV}) were 0.987 and 0.985, and the root mean squared errors in cross-validation (RMSECV) were 0.015 and 0.014, respectively. After, six groups of eight feature wavelengths were chosen from R_S, A_S and E_S based on wavelength selection methods including β-coefficient (βC) of PLSR and the first derivative and mean centering iteration algorithm (FMCIA), and were successively used to build simplified models. The acquired FMCIA-R_S-LWPLSR and βC-R_S-LWPLSR models showed better accuracy than other simplified models, with R^2_P of 0.985 and RMSEP of 0.016 for DMC prediction, and R^2_P of 0.983 and RMSEP of 0.015 for SC prediction, respectively. Besides, the optimal models for MLR and PLSR were obtained using FMCIA on the basis of the E_S. After further reducing the number of feature wavelengths, only six wavelengths (1028, 1068, 1135, 1208, 1262 and 1460 nm) were selected and utilized to develop the simplest FMCIA-Es-MLR model for determining DMC and FMCIA-Es-PLSR model for detecting SC, yielding a reasonable level of accuracy with R^2_P of 0.962 and 0.963 as well as RMSEP of 0.025 and 0.023, respectively. Furthermore, the time series variations of DMC and SC on tuber samples were visualized based on an equation to apply the simplest models to the spectral images.

1. Introduction

Tuber crops are popular staple foods which are consumed as nutritious storage organs in many countries such as Ireland and China (Wang et al., 2013; Zhu et al., 2015; Su and Sun, 2016b). As asexual reproductions, high-yield tubers in terms of potato (*Solanum* Spp.) and sweet potato (*Ipomoea batatas* L.) contain abundant carbohydrates and mineral nutrients, and have been extensively used as processed food (Pedreschi et al., 2012; Sobukola et al., 2015; Muñoz et al., 2017; Schmidt et al., 2016; Oladejo et al., 2017; T. Zhang et al., 2016; H. Zhang et al., 2016; Isik et al., 2016; Ndangui et al., 2014; Amaral et al., 2016; Tarmizi and Niranjan, 2013). Stem tubers of potato and root tubers of sweet potato provide more edible energy per hectare compared with other staple foods such as cereals and legumes (Oerke and Dehne, 2004). Determination of the dry matter concentration (DMC) of these tubers is of fundamental significance to both tuber industry for monitoring processing efficiency, and researchers involved in nutrition studies and energy metabolism (Al-Khusaibi and Niranjan, 2012; Zhang et al., 2011; Pinhero et al., 2016). As a major chemical constituent, tuber starch concentration (SC) plays a vital role in the food industry (Abegunde et al., 2013; Wu et al., 2016; Aina et al., 2012; Blahovec et al., 2012). The tuber quality in terms of nutritional, cooking, and sensory properties can be easily affected by the molecular and crystalline structures of SC (Lisinska and Leszczynski, 1989). Moreover, the differences in tuber microstructure determine the anisotropic distribution of their internal quality parameters. Specifically, DMC is normally higher in the storage parenchyma between the vascular ring and the cortex, and longitudinally reduces in towards the pith

---

[*] Corresponding author.
E-mail address: dawen.sun@ucd.ie (D.-W. Sun).
URLs: http://www.ucd.ie/refrig, http://www.ucd.ie/sun (D.-W. Sun).

http://dx.doi.org/10.1016/j.compag.2017.06.013
0168-1699/

(Pritchard and Scanlon, 1997). Besides, the nutrients of tubers are easily damaged and the hazardous substances can be generated during high temperature processing (Labuza and Tannenbaum, 1972; Skog and Alexander, 2006). In order to cope with this situation, low temperature baking (LTB) was put forward using the temperatures of 65–90 °C to cook food by a prolonged time (Vaudagna et al., 2002; Su and Sun, 2016c). Based on this process, the food nutrients and sensory quality can be effectively preserved. However, the conventional approaches for measuring the internal quality parameters of tuber just acquire the average reference values of food quality, but still labour intensive, inefficient and destructive in most cases (Selim et al., 2009; Tomlins et al., 2012; Aina et al., 2012). Thus, it is of vital importance to develop an on-line analytical method that allows simultaneous determination and visualization of the food quality for the tuber industry.

The non-uniformity in composition results in physicochemical heterogeneity, which was ignored in conventional measurement. Some researchers have realized the capacity of spectroscopic techniques in terms of visible/infrared spectroscopy (Wang et al., 2002; Shao et al., 2011; Lin et al., 2012; Abdel-Nour et al., 2011; Alexandrakis et al., 2012; Shen et al., 2012; Antonucci et al., 2011; Lu et al., 2011; Woodcock et al., 2008), Raman spectroscopy (Almeida et al., 2010; Liu et al., 2013; Sowoidnich et al., 2010; Günaydın et al., 2010; Scheier et al., 2014; Lee and Herrman, 2016), nuclear magnetic resonance (NMR) spectroscopy (Lodi et al., 2007; Zhang et al., 2013; Wu et al., 2014; Shao and Li, 2012; Rodríguez et al., 2013; Ko et al., 2013) and spectral imaging (Shahin et al., 2014; Zhang et al., 2017; Su and Sun, 2017; Ma et al., 2017; Cheng et al., 2015; Zhu et al., 2014), and applied them as fast and non-destructive methods to determine the range of the chemical parameters of staple foods (Helgerud et al., 2015; Flores-Morales et al., 2012). It is worth noting that these techniques may rapidly detect the characteristics of food quality from the spectral features of the samples, but the spatial variation in chemical property is not fully considered. It is of importance to visualize the chemical heterogeneity in determining food quality, because the spatial distribution can exhibit the detailed feature and location of food quality parameter (ElMasry and Nakauchi, 2015). Compared to spectral imaging, the remaining spectroscopic techniques failed to display the spatial information of food chemical properties in a visual pattern especially for heterogeneous foods (Su et al., 2015; Kamruzzaman et al., 2015). As a more intelligent analytical and detective tool that integrated both spectroscopic and computer vision (Jackman et al., 2009; Du and Sun, 2005; Sun and Brosnan, 2003; Jackman et al., 2011) techniques into one system, hyperspectral chemical imaging excelled other techniques in meeting the growing demands of simultaneous acquirement of spatial and spectral information at every location in an image, which has been used for food quality and safety evaluation and analysis (Elmasry et al., 2012; Liu et al., 2014b; Feng and Sun, 2013; Kamruzzaman et al., 2013; Barbin et al., 2013; Feng et al., 2013; Wu and Sun, 2013; ElMasry et al., 2013; Cheng and Sun, 2015b; Xiong et al., 2015; Cheng et al., 2015; Pu et al., 2015; Ma et al., 2016; Cheng et al., 2016a). The acquired spectral image is a three-dimensional data cube with two spatial axes and a wavelength axis. Ordinarily, the spectral axis of multispectral image includes several discontinuous wavelengths at wide intervals while there are hundreds of continuous wavelengths contained in a hyperspectral image (Su and Sun, 2016d; Cen et al., 2014; Qu et al., 2017).

As an accurate and non-invasive determination technique, hyperspectral chemical imaging has a profound and lasting impact on the rapid evaluation of food quality and safety. Based on the information of hyperspectral image and chemical reference values of food quality, the calibration models such as partial least squares regression (PLSR), principal component regression (PCR) and multiple linear regression (MLR) can be built to forecast the corresponding parameters of unknown content (Kamruzzaman et al., 2015; Cheng and Sun, 2015a; Gaston et al., 2010). The applications of this technique have been expanded in various analytical processes (Liu et al., 2014a; Tao and Peng, 2015; Wu et al., 2012). Hyperspectral imaging has currently been tested for mining both the appearance features and internal chemical characteristics of staple food quality such as fungal growth on rice (Siripatrawan and Makino, 2015), hardness of maize kernels (Williams et al., 2009) and sugar content in potatoes (Rady et al., 2015). However, the narrow sampling with hundreds of wavelengths not only increases the memory space size of hyperspectral image but also enhances the correlation between neighbouring bands, which causes data redundancy and multicollinearity.

As the successor of hyperspectral imaging, multispectral chemical imaging system is designed to use several characteristic wavelengths selected from full wavelength range of hyperspectral image. Carrying most of the effective spectral information, characteristic wavelengths are equal or more efficient than full wavelengths (Pu et al., 2015). Multispectral imaging makes non-destructive visual detection more convenient, reliable, and high-efficiency (Cheng et al., 2016a). Moreover, the simplified multispectral imaging system can decrease the costs in manufacture as well. Thus, it is of great importance to select the most effective combination of feature wavelengths. The standard of selection method for characteristic wavelength is to maximize the accuracy of the prediction model using a minimum number of spectral wavelengths. Several wavelength selection approaches including genetic algorithm (GA), β-coefficient (βC) of PLSR, successive projection algorithm (SPA), competitive adaptive reweighted sampling (CARS) have been certified as efficient methods for multispectral imaging and widely applied to selecting characteristic wavelengths for evaluating food quality in the past years (Cheng et al., 2016b; ElMasry et al., 2012; Yu et al., 2014). After the chemical selectivity of spectroscopy is combined with the potency of image visualization, the distribution maps of food quality parameters can be achieved based on the established models using feature wavelengths (ElMasry and Nakauchi, 2016). This is very helpful to inspecting different sources of variations from main phase to single pixel in case of knowing limited foreknowledge of food quality information.

Thus far, no study has been carried out to establish a multispectral chemical imaging technique to measure chemical parameters (DMC and SC) of tubers including potato and sweet potato for staple food industry. It is worthy making a comprehensive research that combines both potato and sweet potato tubers from different regions and extract their characteristic wavelengths to invent on-line spectral imaging systems for certain application. Moreover, these quality parameters are always heterogeneous distribution and dynamically change in tubers during LTB, which makes visual measuring system more indispensable in real-time detection. Additionally, a new approach for wavelength selection should be in the position for designing a more effective multispectral imaging system. Therefore, the focal point of this research was to determine the chemical parameters of these tubers non-destructively and visualize the time series variations of tuber quality parameters rapidly with selected feature wavelengths based on spectral imaging technique. The specific strategies were to: (a) launch an integrated study combining potato and sweet potato tubers to measure DMC and SC using a hyperspectral imaging system in the spectral range of 897–1753 nm, (b) identify the optimal group of feature wavelengths for multispectral imaging by βC of PLSR and first-derivative and mean centering iteration algorithm (FMCIA), (c) establish simplified calibration models to predict DMC and SC using the selected characteristic wavelengths, (d) further reduce the number of feature wavelength for rapid multispectral