Australian Journal of Crop Science AJCS
AJCS 4(3):175-179 (2010) ISSN:1835-2707

Non-destructive measurement of moisture and soluble solids content of Mazafati date fruit by NIR spectroscopy

Seyed Ahmad Mireei*[1], Seyed Saeid Mohtasebi[1], Reza Massudi[2], Shahin Rafiee[1], Atoosa Sadat Arabanian[2], Annachiara Berardinelli[3]

[1]Department of Agricultural Machinery Engineering, Faculty of Agricultural Engineering and Technology, University of Tehran, P.O. Box 4111, Karaj 31587-77871, Iran
[2]Laser and Plasma Research Institute, Shahid Beheshti University, G. C, Evin, Tehran, 1983963113, Iran
[3]Agricultural Economics and Engineering Department, University of Bologna, Piazza G. Goidanich, 60- 47023 Cesena (FC), Italy

***Corresponding author: samireei@ut.ac.ir**

Abstract

The potential of Near-Infrared (NIR) spectroscopy technique was assessed to predict moisture and soluble solids content of Mazafati date fruit. Spectral acquisitions were carried out during four main ripening stages (*Kimri*, *Khalal*, *Rutab* and *Tamr*) using a fast, non-scanning spectrometer in the interactance mode. Two different preprocessing methods; vector normalization and multiplicative scatter correction, and wavelength regions (900- 1700 nm and 1332- 1641 nm) were used to build partial least squares (PLS) regression models. The best predictive models showed coefficient of determination (R^2) values of 0.980 and 0.966 for the moisture and the soluble solids content respectively; the relative residual predictive deviation (RPD) values were 7.1 and 4.8. NIR spectroscopy appeared to be a good method for the assessment of maturity of the analyzed date variety.

Keywords: Near-infrared spectroscopy, moisture content, soluble solid content, date fruit, non-destructive, partial least square method, residual predictive deviation.

Abbreviations: Abs_absorbance; MC_moisture content; MSC_multiplicative scatter correction; NIR_near-infrared; PLS_ partial least squares; r_correlation coefficient; R^2_coefficient of determination; RMSECV_root mean square error of cross validation; RMSEP_root mean square error of prediction; RPD_residual predictive deviation in test set validation; RPDCV_ residual predictive deviation in leave one out cross Validation; SEP_standard error of prediction; SSC_soluble solids content; VN_vector normalization.

Introduction

Date fruit (*Phoneix dactylifera*) is one of the most important and oldest agricultural productions in Iran. Many people, especially in southern regions of Iran, earn their livelihood from date fruit. With production of 1,000,000 tons of date fruit and exportation of 143,351 tons in 2006, Iran is one of the largest producers of dates in the world. There are about 400 different varieties of date fruit in Iran, however, the Mazafati variety is the most famous and delicious one and is usually considered as a soft or wet variety (FAOSTAT, 2006). Some date fruit ripen early in the summer, whilst others are not mature until the end of the season (August/September), and during maturity, physical and chemical properties of date fruits change considerably (Dowson, 1982). Because of the single-pass harvesting method of date fruits, the harvesting should include the date fruits in all ripening periods consisting in four distinct stages of the ripening (Sahari et al., 2007; Imad et al., 1995; Al-Shahib and Marshall, 2003). In the first stage, called *Kimri*, dates are green with a high moisture content and acidity and a low sugar content. In the second stage, called *Kharak* or *Khalal*, the skin of Mazafati dates becomes red, the acidity and the size of the dates decrease while the sugar increases. At the *Rutab* stage, the Mazafati dates become softer and brown with a higher sugar content, and at the final stage, called *Tamr*, dates are black with a low moisture content and high sugar content and have good storability properties (Al-Shahib and Marshall, 2003). After fruit harvesting, dates can be stored in cooled conditions before consumption. During storage, the unripe dates will ripen properly according to the amount of sugar absorbed from the tree. The moisture and the sugar content are the main factors affecting the ripening of dates (Schmilovitch et al., 1999). Measurement of moisture content (MC) and soluble solids content (SSC) of date fruit using common destructive methods is a very time-consuming task. Generally, determination of SSC of date fruit is difficult because there is no extractable juice

Table 1. Statistical parameters of date fruits for both calibration and validation data sets

Parameter	Data Set	No. of Samples	Range	Mean	Standard Deviation	Coefficient of Variability(%)
MC (%)	Calibration/leave one out CV	120	42.31-86.44	67.14	13.79	20.44
	Test set validation	40	46.20-85.71	66.28	13.58	20.49
SSC (°Brix)	Calibration/leave one out CV	120	9.2-44.0	25.0	10.7	42.9
	Test set validation	40	10.4-44.4	25.9	10.6	40.8

even in the *Kimri* stage when moisture content is more than 80%. At present, near-infrared (NIR) spectroscopy represents a non-destructive, rapid and accurate method for measurement of MC and SSC in several varieties of fruits (Nicolaï et al., 2007). Using the complete range of near-infrared wavelengths (800-2500 nm) in the reflectance mode, several studies were conducted to predict these parameters for apples (Liu and Ying 2005; Yan-de et al., 2007) and citrus fruit (Huishan et al., 2005). Other interesting studies were carried out using reduced NIR wavelength ranges for the prediction of the SSC in pears (Chen et al., 1999), cherries (Lu, 2001), apples varieties (Lu and Ariana, 2002 and Alamar et al., 2007) and for assessment of MC in avocado fruit (Clark et al., 2003) and potatoes (Singh et al., 2004). Though different acquisition methods and data analysis were used, all of these studies verified the NIR spectroscopy power for assessment of MC and SSC. Few studies have been done on date fruit regarding MC and SSC. Schmilovitch et al., (1999) used a semi-automatic system based on NIR spectroscopy in the range of 1200- 2400 nm to predict the SSC and MC of Hayani dates; the results showed a standard error of prediction (SEP) of 1% for both SSC and MC and a correlation coefficient (r) of 0.9. The objective of the present study was to develop a non-destructive rapid method based on NIR spectroscopy to determine the MC and SSC of Mazafati date fruits during maturation. This study could represent a very important tool to make decisions about the optimum harvesting time.

Materials and methods

Samples

Measurements were carried out on 160 Mazafati date fruits, harvested in two different periods (August 1, 2008 and September 1, 2008), from the Jiroft area (Kerman province, Iran). After each harvesting, fruits were stored at 2°C for 10 days at most. Date fruits were characterized by the following ripening stages: *Kimri* (25%), *Khalal* (38%), *Rutab* (25%) and *Tamr* (12%) of total samples. These ripening stages are representatives of the entire Mazafati maturity period (Fig. 1). Before spectral acquisition and in order to have a steady state of fruit temperature, fruits were equilibrated for about 6 hours to room temperature (about 22 °C).

Spectral acquisitions

NIR spectra were acquired using an EPP2000 NIR InGaAs Spectrometer (StellarNet, Inc. Oldsmar, Florida, USA) in the 900–1700 nm wavelength range (2.5 nm of resolution) (SpectraWiz software, StellarNet, Inc). The spectrometer, a non-scanning portable instrument, equipped with an InGaAs detector with 1024 pixels (25µm by 500µm tall), was connected to a computer via high speed USB 2 port. Light from the tungsten halogen

Fig 1. Mazafati date fruits in different ripening stages

light source (SL1-CAL Tungsten Halogen Light Source, StellarNet, Inc.) was carried to the sample through six optic fibers in a bifurcated and interactance fiber optic probe (R400-7-VISNIR, StellarNet, Inc.), diffused through the date flesh and reflected back through the central optic fiber to the spectrometer (Fig. 2). Each optic fiber in the interactance probe (88.9 mm length and 6.35 mm outer diameter) is 400 µm thick. A reflected spectrum was acquired with an integration time of 240 milliseconds (10 scans per sample) and active temperature compensation mode (every 15 seconds). A dark scan was performed at the beginning of each experiment session and a reference measurement was carried out using a 50 mm diameter white reflectance standard made of Halon at the beginning of each experiment session and after each twenty sample measurements. This white standard reflects more than 97% of the light from 900-1700 nm. The spectrometer software automatically divided the difference between the date reflectance spectra and the dark spectrum by the difference between the reference and dark spectrum to provide the relative reflectance ratio. Dates were placed centrally upon the hole, with stem–calyx axis horizontal and a distance of about 5 mm from the probe (Fig. 2). Three spectra were acquired at about 120° intervals around the date equatorial axis. For the dates which were transferring from one stage to another one, these three spectra included the ripe portion, the unripe portion and the space between the ripe and unripe portions as shown in Fig. 3. For each date, the three acquisitions were averaged and the mean spectrum was calculated.

Reference Measurements

Immediately after spectral acquisitions, date fruits were cut in the length in two halves and the inside seeds removed. One half was used to measure MC and the other half to assess the SSC. MC was determined by using the method described by Keramat Jahromi et al., (2008). For SSC, the half fruit was mixed with distilled water in the mass ratio of three and then filtered and finally the filtrate was used for SSC determination using

Table 2. PLS model characteristics for the prediction of date fruit moisture content (MC)

preprocessing	Wavelength range	Calibration			Leave one out cross validation			Test set validation			
		LV	R^2	RMSEC	R^2	RMSECV	RPDCV	LV	R^2	RMSEP	RPD
VN	1332-1641	7	0.991	1.37	0.989	1.48	9.32	6	0.980	1.91	7.11
	926-1700	9	0.991	1.42	0.987	1.61	8.48	6	0.978	1.99	6.82
MSC	1332-1641	6	0.991	1.34	0.989	1.47	9.38	5	0.977	2.02	6.72
	926-1700	8	0.990	1.54	0.986	1.67	8.18	6	0.979	1.92	7.07

LV: number of latent variables.
VN: vector normalization.
MSC: multiplicative scatter correction.
RPDCV: residual predictive deviation in leave one out cross validation.
RPD: residual predictive deviation in test set validation.

a digital refractometer (Abbe refractometer, Sun Instruments Corp., Torrance, CA, USA). The real °Brix value was then calculated by multiplying the °Brix value of mixture to the ratio of the weight of the diluted mixture and the date (AOAC official method for analyzing, 1998).

Data preprocessing and analysis

All absorbance spectra (abs=log[1/R], R: reflectance) were analyzed using OPUS software (ver. 5.5, Bruker Optik GmbH). Two different data preprocessing techniques were considered: vector normalization (VN) and full multiplicative scatter correction (MSC). These techniques are commonly used to eliminate the irrelevant information from spectra due to unknown sources such as surface irregularities, distance variation of sample and detector (Lu, 2001). In particular, VN normalizes a spectrum by subtracting an average intensity value from it and MSC is able to correct multiplicative effects due to scatter by performing a linear transformation of each spectrum. Partial least square regressions (PLS) were used in order to build predictive models of MC and SSC for Mazafati date fruits. These models were evaluated in terms of R^2 and root mean square error of cross validation (RMSECV) and prediction (RMSEP) (test set validation). In order to avoid under or over fitting, the optimum number of latent variables in PLS models was determined by finding the minimum plot of the RMSECV of the parameter in question versus the number of latent variables (Martens and Naes, 2001). 75% of the data set (calibration set) was used for leave one out cross validation and the remaining 25% for test set validation. The accuracy of the regression models was assessed by calculating the residual predictive deviation values by dividing the standard deviation SD of the reference values by the standard error of cross validation or prediction (Williams and Norris, 2001). In order to improve the predictive power of the models and reduce the acquisition time, a narrow wavelength range was optimized with reference to the best obtainable R^2 and error values for MC and SSC.

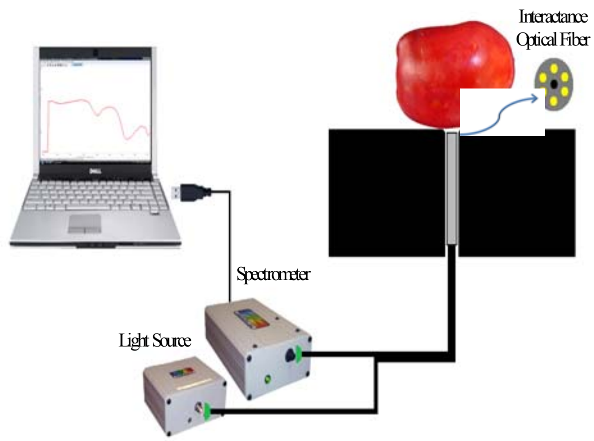

Fig 2. A schematic of the NIR instrument set up

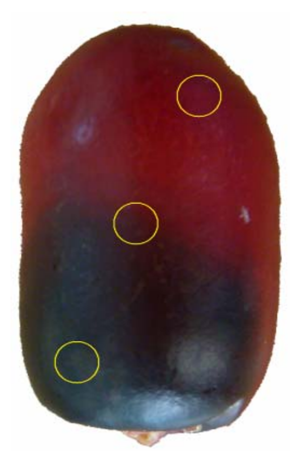

Fig 3. Scanning places for fruits which were transferring from one ripening stage to the next

Results and Discussion

Fig. 4 shows the means values of the absorbance of Mazafati dates in the *Kimri*, *Khalal*, *Rutab* and *Tamr* ripening stages. Absorption spectra have a very broad peak around 970 nm which is related to the second -OH overtone of water and relatively broad peaks around 1190 nm and 1462 nm which are related to –CH stretch first overtone and the first –OH overtone of water respectively. Table 1 summarizes the statistical parameters of the fruits used for calibration and validation of date fruit samples. In both calibration and test sets, the variability of MC (20.44% and 20.49% for calibration set and test set, respectively) and SSC (42.9% and 40.8% for calibration set and test set, respectively) are close to each other, thus MC and SSC have a good variability through sample sets.

Table 3. PLS model characteristics for the prediction of date fruit soluble solids content (SSC)

preprocessing	Wavelength range	Calibration			Leave one out cross validation			Test set validation			
		LV	R^2	RMSEC	R^2	RMSECV	RPDCV	LV	R^2	RMSEP	RPD
VN	1332-1641	6	0.967	2.02	0.962	2.08	5.14	4	0.951	2.36	4.49
	900-1700	8	0.966	2.06	0.956	2.26	4.73	7	0.955	2.24	4.73
MSC	1332-1641	6	0.966	2.03	0.962	2.09	5.12	5	0.948	2.42	4.38
	900-1700	8	0.971	1.89	0.962	2.11	5.07	6	0.966	2.21	4.80

LV: number of latent variables.
VN: vector normalization.
MSC: multiplicative scatter correction.
RPDCV: residual predictive deviation in leave one out cross validation.
RPD: residual predictive deviation in test set validation.

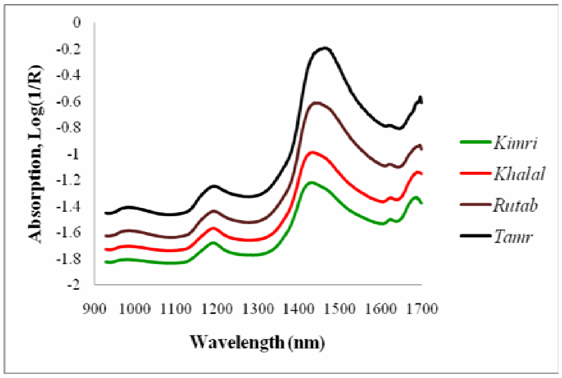

Fig 4. Raw absorption spectra (mean value) of Mazafati dates in Kimri, Khalal, Rutab and Tamr ripening stages

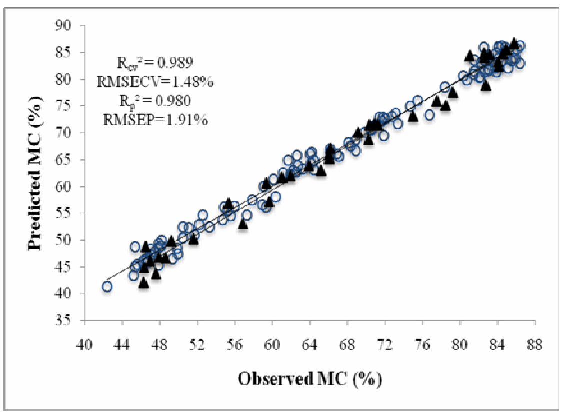

Fig 5. MC prediction resulted from leave one out cross validation (○) and test set validation (▲) using vector normalization of spectra in the wavelength range 1332-1641

Moisture content prediction (MC)

Table 2 shows the results for the PLS models set up for the MC prediction and for both VN and MSC preprocessing methods. The reported parameters were obtained in the full spectral range 900-1700 nm and in the narrow range 1332-1641 nm. PLS model validations showed R^2 values ranging from 0.989 (RMSECV, 1.47%) to 0.986 (RMSECV, 1.67%) and from 0.977 (RMSEP, 2.02%) to 0.980 (RMSEP, 1.91%) for the leave one out cross validation and for the test set validation, respectively. The obtained RPD values were not lower than 6.72; according to Williams and Norris (2001), RPD values ranging between 6.5-8.0 indicate a very good performance of the model, while the values higher than 8.1 indicate an excellent prediction of composition. The

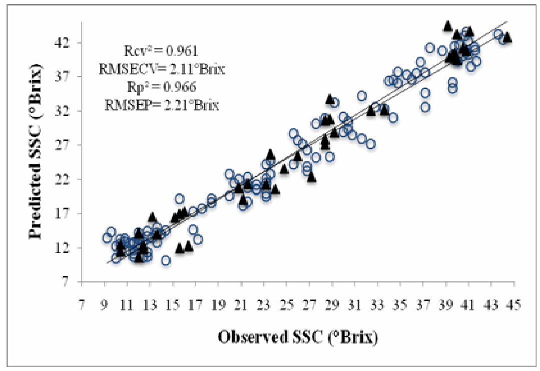

Fig 6. SSC prediction resulted from leave one out cross validation (○) and test set validation (▲) using multiplicative scatter correction in full wavelength range 900-1700 nm

model performance varied slightly between the two wavelength ranges and preprocessing methods. The best results were obtained using the narrow range (1332-1641 nm) and, in general, the VN preprocessing method gave the best predictions. This may be due to the strong absorption bands of pure water at 1460 nm. Fig. 5 shows the values predicted by the model versus the observed values of MC obtained by performing the leave one out cross validation and the test set validation. The optimum number of latent variables for this model was selected as 7.

Soluble solids content prediction (SSC)

The results of the PLS regressions performed for the SSC for both VN and MSC preprocessing methods are summarized in Table 3. As stated for the MC prediction, the reported parameters were obtained in the full spectral range 900-1700 nm and in the narrow range 1332-1641 nm. PLS model validations showed R^2 values ranging from 0.956 (RMSECV= 2.26 °Brix) to 0.962 (RMSECV= 2.08 °Brix) and from 0.948 (RMSEP= 2.42 °Brix) to 0.966 (RMSEP= 2.21 °Brix) for the leave one out cross validation and for the test set validation, respectively. Looking at the RPD values (not lower than 4.38), these models could be used for quality control (Williams and Norris, 2001). As for MC, no great differences emerged between the predictive power of the models built using different wavelength ranges and preprocessing methods. In terms of prediction, the highest R^2 values were obtained with the full range of acquisition and the MSC method. The values predicted by the model versus the observed values of SSC obtained by performing the leave-one out cross validation and the test set validation are shown in Fig. 6. The number of 8 latent variables was selected for this PLS model.

Conclusions

The obtained results showed that NIR spectroscopy in the interactance mode combined with the PLS regression is a good and reliable technique to measure the maturity indices of Mazafati date fruits. The moisture content (MC) and the soluble solids content (SSC) can be predicted with root mean square errors (RMSEP) up to 1.91% (R^2= 0.980) and 2.21 °Brix (R^2= 0.966), respectively. For both parameters, the model predictive power was not greatly influenced by the preprocessing method used (vector normalization, VN, or multiplicative scatter correction, MSC) and by two wavelength ranges (900-1700 nm and 1332-1641 nm). The high residual predictive deviation (RPD) values (up to 7.1 and 4.8 for the moisture and the soluble solids content, respectively) suggest that these models could be used to accurately assess the quality of Mazafati date fruits during the maturity period.

Acknowledgments

The authors would like to thank Research Deputy of University of Tehran for its financial support.

References

Alamar MC, Bobelyn E, Lammertyn J, Nicolaï B.M, Moltó E (2007) Calibration transfer between NIR diode array and FT-NIR spectrophotometers for measuring the soluble solids contents of apple. Postharvest Biology and Technology 45: 38-45.

Al-Shahib W, Marshall RJ (2003) The fruit of the date palm: its possible use as the best food for the future?. International Journal of Food Sciences and Nutrition 54 (4): 247-259.

Association of Official Agricultural Chemists (1998) Official Methods of Analysis. A.O.A.C. 16th Ed. Published by A.O.A.C. Washington, D.C. (U.S.A).

Chen C, Shaw J (1999) Determination of the sugar content and acidity of pears by a portable near-infrared spectrophotometer. Journal of Agricultural Machinery 8(1): 49-57.

Clark CJ, McGlone VA, Requejo C, White A, Woolf AB (2003) Dry matter determination in Hass avocado by NIR spectroscopy. Postharvest Biology and Technology 29: 300-307.

Dowson V HW (1982) Date Production and Protection. Plant Production and Protection Paper No.35. 1982, Rome, Italy, FAO.

FAOSTAT (2006) Statistical Year Book of FAO, Available in: http:// faostat.fao.org.

Huishan L, Yibin Y, Huanyu J, Yande L, Xiaping F, Wang J, Application Fourier Transform Near Infrared Spectrometer in Rapid Estimation of Soluble Solids Content of Intact Citrus Fruits. ASAE Paper Number: 053042. St. Joseph, Mich.: ASAE.

Imad AA, Abdul Wahab KA (1995) Chemical composition of date varieties as influenced by the stage of ripening. Journal of Food Chemistry 54: 305-309.

Keramat Jahromi M, Jafari A, Rafiee Sh, Mirasheh R, Mohtasebi SS (2008) Changes in physical properties of date fruit (cv. Shahani) during three edible stages of ripening. American-Eurasian Journal of Agricultural and Environmental Science 3(1): 132-136.

Liu Y, Ying Y (2005) Use of FT-NIR spectrometry in non-invasive measurements of internal quality of Fuji apples. Journal of Postharvest Biology and Technology 37: 65-71.

Lu R (2001) Predicting Firmness and sugar content of sweet cherries using near-infrared diffuse reflectance spectroscopy. Transactions of the ASAE 44 (5): 1265-1271.

Lu R, Ariana D (2002) A near-infrared sensing technique for measuring internal quality of apple fruit. Transactions of the ASAE 18(5): 585-590.

Martens H, Naes T (2001) Near-Infrared technology in the Agricultural and Food industry. St. Paul, MN: American Association of Cereal Chemists, Inc.

Nicolaï BM, Beullens K, Bobelyn E, Peirs A, Saeys W, Theron KI, Lammertyn J (2007) Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. Journal of Postharvest Biology and Technology 46: 99-118.

Sahari MA, Barzegar M, Radfar R (2007) Effect of Varieties on the Composition of Dates (Phoenix dactylifera L.). Journal of Food Science and Technology International 13(4): 269-275.

Schmilovitch Z, Hoffman A, Egozi H, Ben-Zvi R, Bernstein Z, Alchanatis V (1999) Maturity determination of fresh dates by near infrared spectrometry. Journal of the Science of Food and Agriculture (79): 86-90.

Singh B, Wang N, Prasher S, Ngadi M, A Spectroscopic Technique for Water Content Determination in Potato. ASAE Paper Number: 046116. St. Joseph, Mich.: ASAE.

Williams PC, Norris K (2001) Near-Infrared technology in the Agricultural and Food industry. St. Paul, MN: American Association of Cereal Chemists, Inc.

Yan-de L, Yi-bin Y, Xiaping F, Huishan L (2007) Experiments on predicting sugar content in apples by FT-NIR Technique. Journal of Food Engineering. 80: 986-989.