

NON DESTRUCTIVE VISIBLE AND NIR SPECTROSCOPY MEASUREMENT FOR THE DETERMINATION OF APPLE INTERNAL QUALITY

Moons, E.¹ and Sinnaeve, G.²

CRA, Département Qualité des Productions Agricoles,
Chaussée de Namur, 24
5030 Gembloux
Belgium

Tel: +32/81/61 16 62 Fax: +32/81/61 58 78 E-Mail: moons@cragx.fgov.be¹; sinnaeve@cragx.fgov.be²

Dardenne, P.

CRA, Département des productions animales
rue de Serpont, 100
6800 Libramont
Belgium

Tel: +32/61/23 10 10 Fax: +32/61/23 10 28

Keywords: acidity; apple grading; Brix; diode array; near infrared; non destructive quality control

Abstract

A non destructive automatic near infrared sensor would be greatly appreciated among package storage houses and great store market for grading apple according to their taste quality. For this purpose, a NIR spectrometer DA 7000 (Perten), a polychromatic spectrometer fitted with InGaAS diode array detectors has been tested. The measurement time and prediction only take 1 second and could even be lowered with the Perten on line software. The non destructive reflectance spectra are taken in the four equidistant locations along the equator on the whole apples. A modified partial least square regression method based on the NIR spectra (math treatment 2,4,4,1) and lab values, gave good results. Determination coefficients and standard errors of prediction for 300 apples were respectively : 0.86-0.45 for brix, 0.88 - 1.33 ml of NaOH 0.1N for titrable acidity, 0.86 - 0.56% for dry matter, 0.78 - 0.10 for pH, 0.64 - 0.74 kg/cm² for firmness. A high correlation between the second derivative of spectra and diameter was also found : 0.87-0.17 cm.

1. Introduction

More and more consumers' requests have been noticed recently to buy a prime quality fruit with a better taste upon higher rewarding price. Three major parameters have to be taken into account to determine the internal quality and the taste of apple. These are hardness, sugar content and titrable acidity.

At the present time, grading of apples is based only on the external aspect. Fruits are sorted visually and manually according to their size and colour. Non marketable apples with surface defects such as scabs, breaks and bruises are discarded.

The goal of our study is to develop a grading machine which is able to classify apples into different classes according to their supposed organoleptic quality.

Non-destructive NIR methods have already been used to study dry matter content in onions (Birth *et al.*, 1985), potatoes (Dull *et al.*, 1989) and whole dades (Dull *et al.*, 1991), internal quality in peaches, nectarines (Slaughter, 1995), raisins (Huxsoll *et al.*, 1995) and in pineapple (Guthrie *et al.*, 1996).

Davenel *et al.* (1988) predicted sugar concentration of 320 apples by near infrared techniques with a standard error of prediction of 6.8 g/l of sugar and a correlation coefficient of 0.96. Bellon *et al.* (1993) developed a NIR instrument coupled with optical fibres to detect sugar at the speed of three apples per second. Standard error of prediction was 2.4 g/l of glucose.

Kawano (1994) described a Multi Purpose Sensor (MPS) used by Mitsui Mining and Smelting Co., LTD., for grading apples, pears and peaches. Three apples were sorted per second.

2. Materials and methods

2.1. Instrument

The NIR spectrometer used is a DA 7000 from Perten Instruments. This post-dispersive polychromator generates spectra from 400 to 1700 nm with a spectral resolution of 5 nm. The sample is measured in the reflectance mode as shown in figure 1. A polychromatic beam falls on the sample. Some light enters, undergoes multiple scattering and exits the sample. Such a radiation is said to be diffusely reflected. This reflected light is sent to the grating, split into wavelengths and then collected by diode array detectors. These detectors are made in Indium Gallium Arsenite for the NIR range (1100 to 1700 nm) and in Silicon for the visible and the beginning of near infrared range (400 to 1098 nm). The fruits are just laid down on a static ring at 20 cm below the instrument. The spectrum measurement takes 1 second.

2.2. Apples

The apples were measured during two years (1996 and 1997) just before harvest, at harvest and during storage time. The apples were also selected to cover the entire variability of studied parameters (brix, acidity and hardness). Apples of the main varieties found in Belgium were selected: Jonagold, Golden delicious, Belle de Boskoop and Jonagored. Boskoop is interesting to include in the data set because it has very high values of acidity and hardness.

2.3. Reflectance measurements

Whole apples were measured without any destruction nor damage to the fruit. The reflection spectra were taken in four equidistant positions along the equator of the apple in order to obtain a good estimation of the whole fruit and to eliminate spatial variability. Each absorption spectrum used in the calibrations was the average of the four spectra obtained for each apple. A Spectralon® plate was used as the external optical standard reference, before each series of measurement. The collected spectral data were processed by Grams 32 and Infrasoft International software (ISIV4).

2.4. Wet chemical analyses and hardness

The hardness is measured by a texture meter (QTS) which monitors the force to compress the apple flesh with a probe of 1 cm² surface area at the speed of 100 mm/min. The results are expressed in kg/cm².

A manual Atago refractometer (Atago N 20) is used to measure total soluble content (brix value) of juice extracted from apples. The accuracy of the method is 0.2 brix.

Apple juice acidity is obtained by titration using a NaOH 0.1N solution. Results are expressed in ml of NaOH needed to titrate 10 ml of apple juice until the pH's solution reaches 8.2 value.

pH was systematically measured before titrating.

For the dry matter determination, drying in vacuum oven at 70°C during 20 hours is used as standard method.

Physical parameters such as size are also measured because the rather large area illuminated by the instrument allows to discriminate the apple from the back ground.

2.5. Principel Component Analysis and partial least square regression

Principal component analysis (PCA) is a mathematical procedure that reduces the number of variables X (wavelengths in the case of NIRS) to a new system of axis which accounts for the largest possible amount of variance data as possible. Each principal component is a set of loadings (weights), one for each wavelength. A score is computed by multiplying spectral absorbance values by these weights and by summing the results. The aim of PCA is to detect outlier spectra and to find links between different group criteria such as years of harvest, varieties, etc.

2.6. Multilinear regression and partial least square regression

The acquisition of NIR spectra for a set of calibration (called calibration set) allows to establish calibration models relating spectral data to chemical reference values for each component to analyse. Such calibration models can be checked by using a set of validation samples (called validation set) that are no part of calibration set, but are similar to those respect to constituent range, physical state, origin, etc. The check proceeds as follows: 1) spectral data are collected and interpolated into a calibration model to predict concentration values; 2) these estimates are then compared with chemical reference values. Two calibration methods are used in this paper: Multilinear regression (MLR) and Partial least squares regression (PLS). The first one is the simplest statistical method of predicting one variable from one or more other variables. The mathematical model fit by multiple regression is the following:

$$Y = B_0 + B_1X_1 + B_2X_2 + \dots + B_pX_p$$

Y is the variable to be predicted

B₀ is the equation intercept

B₁ through B_p are the scaling factors that relate changes in the X variables to Y

P is the number of X variables

The performance of a calibration model to predict component concentrations based on spectral data can be assessed by using the standard error of prediction and the determination coefficient between estimates and chemical reference values. The first one should be lower than standard deviation of the data set and allows to make quantitative prediction when the ratio of SD/SEP is equal or superior to 3. The higher this ratio is, the better is the model.

- SEP: Standard Error of Prediction. This parameter is usually estimated on prediction data set.

$$SEP = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n}}$$

y_i laboratory value

\hat{y}_i predicted value

n: number of samples

p: number of terms in the equation

$$R^2 = 1 - \frac{SEC_r}{SEC_t}$$

SEC_r is the sum of square of residual differences between lab and predicted values

SEC_t is the total sum of square of the differences of the variable to explain

The second one, the Partial least square regression method uses the full spectrum.

It is similar to PCA method but differs that the loadings calculations include laboratory values. The modification of the classical PLS algorithm involves the standardization of the residuals after each iteration. The calibration routine runs automatically cross validations (4 groups) and gives standard error (SEVC) and determination coefficient (R^2V) of prediction. For more explanation, see Shenk et Westerhaus (1995).

3. Results and discussion

3.1. Apple set

A set of 300 apples was used for the calibration. The number of samples (N), mean, standard deviation (std deviation), minimum and maximum levels (range) of physical and chemical parameters measured by the laboratory methods are given in table 1.

Table 2 explains some maturity evolutions observed in an apple : when sugar increases with maturity, the acidity and hardness parameters (except pH) decrease. This table describes a general trend between the four parameters, but all these evolutions are dependent on the varieties and storage conditions.

Figure 2 shows spectral variations of the different apple varieties in the PCA space. The set of samples was mainly composed by the Jonagolds. This explains the rather large area covered by this variety. The Jonagored area (mutant variety from Jonagold) overlaps the Jonagold's one. Golden and Boskoop spectra occupy distinct specific areas.

3.2. Data mathematical treatments and calibrations

The highest correlations between the spectra and the reference values were found by using the second derivative of spectra ($d^2\log I/R$). This method reduces spectral variations due to radiation scattering and sample size.

A first calibration using a MLR (multiple linear regression) method has been carried out. In this case the term SECV is not an actual validation on independent sets as with the PLS method, but SECV by MLR is a root mean square (RMS) of the Studentized residuals (Gunst and Mason, 1980). The calibration results obtained are shown in table 3. High RSQV (>0.80) can be obtained with brix, acidity, dry matter and diameter and allows to grade apples into different classes of quality.

A second calibration using a Modified Partial Least Square regression method has been loaded on the same data set. The results (table 4), achieved by this method are more accurate (lower SECV) except for dry matter and hardness parameters. These ones give similar results. The results obtained for brix, acidity, dry matter and diameter parameters allow quantitative predictions. For brix and dry matter ones the NIR range (1100-1700 nm) was sufficient.

3.3. Effect of one apple variety on the accuracy of prediction

A calibration was developed only for the Jonagold variety to check the performance of specific equations compared with those obtained with the whole data set. Jonagold was chosen as it is the main variety cultivated in Belgium (65% of the Belgian apple production) and for this variety we collected 200 spectra. Chemical and physical reference values are given in table 5 and calibrations results in table 6. The brix and hardness ranges are tighter; this is due to the lower variability inside one variety.

The results show an improvement in accuracy for acidity and hardness (lower SEC and SECV). For diameter parameter the range was not large enough to make a good calibration. The SEC 0.32 obtained for brix parameter is equivalent to an error of prediction of 2.83 g/l of sugar.

3.4. Repeatability between duplicate NIR measurements

A small set of 11 apples was measured in duplicate to have an estimate of the repeatability. Each apple was placed twice in the same position for measurement. The standard errors of differences (SEDr) were calculated from the predictions obtained for the 2 spectra on the same spot of the apple. The results are reported on table 7.

Thus the repeatability of DA7000 is measured by SEDr. This SEDr value is much lower for brix and pH than their respective SEC on table 3. It means that it is not necessary to measure more than once the same position, but the variability of the predictions at the each positions on the apple are much higher. This observation shows that apples are not homogeneous. This was confirmed by several brix and hardness lab measurements made on the same apple (see table 8). So the NIR prediction from one measure includes the non precision of the laboratory method and the non homogeneity of apple. If the number of measures around the fruit is increased, the SEP decreases. Actually the variability due to the inhomogeneity of the apple decreases.

4. Conclusions

The results we obtained in this study show that several parameters of intact apples can be measured by NIR with good accuracy even with a limited wavelengths range (400-1700 nm). The best results were achieved by using a Modified Partial Least Square (MPLS) regression method on the second derivative of the spectra. The statistical performances of the predictive models are : $R^2V = 0.86$, $SECV = 0.45$ for brix , $R^2V = 0.88$, $SECV = 1.33$ for acidity ; $R^2V = 0.64$, $SEP = 0.74$ for firmness.

Graders could add NIR techniques to the visual grading to improve the quality control of their products. In practice, apples are sold in batch and the internal quality of the whole batch is estimated by the average of the lab values of sample sub set. With this method all apples could be controlled.

The multiple linear regression allows to choose the wavelengths which correlates best with apple internal quality parameters and cheaper NIR instruments with adequate filters corresponding to these wavelengths could be used.

Acknowledgements

This work was made possible through the financial support of the Belgian Agricultural Ministry, research and development DG6.

References

- Bellon V., Vigneau, J.L., Leclercq, M. 1993. Feasibility and performance of a new multiplexed, fast low cost fiber optic NIR spectrometer for the on-line measurement of sugar in fruits. *Applied spectroscopy*, 47: 1079-1083
- Birth, G.S., Dull, G.G., Renfro, W.T., Kays, S.J. 1985. Non destructive spectrophotometric determination of dry matter in onions. *Journal of the American Society for Horticultural Science*, 110(2): 297-303
- Davenel, A., Crochon, M., Pourcin, J., Verlaque, P., Devaux, M.C. 1988. On line infrared measurements in food processing . Presentation at agricultural Engineering International conference
- Dull, G.G., Birth, G.S., Leffler, R.G., Zaltzman, A., Schmilovitch, Z. 1991. The near infrared determination of moisture in the whole date. *HortScience*, 26(10): 1303-1305
- Dull, G.G., Birth, G.S., Leffler, R.G. 1989. Use of near infrared Analysis for non destructive measurements of dry matter in potatoes. *Amer.Potato J.*, 66: 215-225
- Gunst, F., Mason, L. 1980. Regression analysis and application. Marcel Dekker inc: p400, USA

- Huxoll, C.C., Bolin, H.R., Mackey, B.E. 1995. Near infrared analysis potential for grading raisin quality and moisture. *Journal of food sciences*, 60 (1): 176-180
- Kawano, S. 1994. Non destructive NIR quality evaluation of fruits and vegetables in Japan. *NIRnews*, 5 (6): 10-12
- Slaughter, D.C. 1995. Non destructive determination of internal quality in peaches and nectarine. *Transation of the Asae*, 38 (2): 617-623

Tables

1. Chemical and physical reference values of a 300 apple set

	N.	Range	Mean	std Deviation	V
%Brix	290	9.4-19.4	13.5	1.39	10
pH	228	3.07-4.11	3.49	0.23	6.5
Acidity	230	2.83-21.99	8.99	4.39	49
Hardness	291	3.28-11.02	7.12	2.01	18
%Dry Matter	136	10.8-24.07	16.4	1.96	12
Diameter (cm)	183	6.89 -10.2	8	0.61	8

2. Correlation coefficients between the different parameters of quality

	pH	Acidity	Hardness
Brix	0.03	-0.45***	-0.44***
pH		-0.56***	-0.47***
Acidity			0.79***

*** : P<0.001% Probability to be significantly different from 0.

3. Calibration results for quality parameters of the 300 apple set with MLR method

	SEC	RSQ	SECV	RSQV	MLRT
Brix	0.46	0.87	0.47	0.86	9
pH	0.11	0.76	0.11	0.75	10
Acidity	1.56	0.86	1.60	0.86	7
Hardness	0.70	0.67	0.73	0.64	8
Dry Matter	0.64	0.86	0.67	0.85	8
Diam	0.14	0.95	0.14	0.98	4

SEC: residual standard deviation of calibration

RSQ: determination coefficient of calibration

MLRT : number of wavelengths in the stepwise calibration

4. Calibration results for the determination of quality parameters with modified PLS method

	SEC	RSQ	SECV	RSQV	PLST
Brix	0.40	0.89	0.45	0.86	10
pH	0.09	0.82	0.10	0.78	10
Acidity	1.19	0.91	1.33	0.88	12
Hardness	0.71	0.66	0.74	0.64	7
Dry mat	0.53	0.89	0.69	0.82	9
DIAM	0.13	0.92	0.17	0.87	10

SECV: standard error of cross validation (4 groups)

RSQV: determination of cross validation

PLST : number of PLS terms

N: number of samples

5. Chemical and physical reference values of the Jonagold variety data set

	N	Range	Mean	S.D.
Brix	202	9.80 -16.70	12.12	1.53
pH	195	3.21-4.11	3.29	0.19
Acidity	202	3.6 - 15.02	10.26	3.89
Dry mat	161	13.4 - 19.61	16.12	1.60
Hardness	144	3.86 -11.02	9.93	1.28
Diameter	185	6.90 -9.02	7.77	0.36

6. Calibration results for the determination of quality parameters with modified PLS method of Jonagold data set.

	SEC	RSQ	SECV	RSQV	PLST
Brix	0.32	0.93	0.46	0.86	8
pH	0.13	0.71	0.14	0.67	10
Acidity	0.95	0.95	1.24	0.92	7
Dry mat	0.75	0.78	0.87	0.69	7
Hardness	0.60	0.64	0.62	0.62	8

7. Repeatability between duplicates obtained on the same spot

	Brix	pH	Acidity	Hardness
SEDr	0.07	0.003	0.27	0.36

8. Variability of 8 brix and 4 hardness lab measurement inside the apples

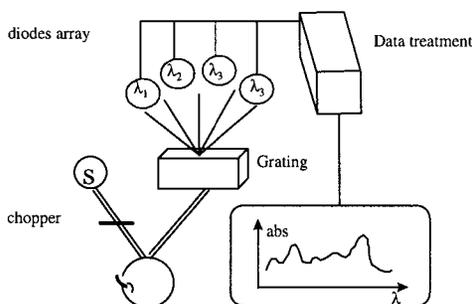
Jonagold	N	Average	SD	min	max
brix	30	12.97 brix	0.34	10.8	17.0
Hardness	40	5.60 kg/cm ²	0.38	3.6	9.4

9. Variation of the standard error of prediction of several fruit parameters according to the number of scans around apple

	1 scan	2 scans	3 scans	4 scans
<i>Brix</i>	0.95	0.73	0.58	0.46
<i>pH</i>	0.14	0.13	0.12	0.10
<i>Acidity</i>	2.37	1.67	1.50	1.36
<i>FSP</i>	1.02	0.94	0.91	0.78
<i>DIAM</i>	0.28	0.26	1.08	0.23
<i>MS</i>	1.52	1.28	1.23	0.99

Figures

1. Schematic view and picture of the post-dispersive diode array instrument



2. Variability of 300 apple spectral data in the two first axis of the PCA

