Near-infrared spectroscopy for non-destructive prediction of maturity and eating quality of 'Carabao' mango (*Mangifera indica L.*) fruit

Yvonne Q. Polinar[1], Kevin F. Yaptenco[2], Engelbert K. Peralta[3], Josephine U. Agravante[4]

(1. *Cebu Technological University – Barili , Cebu, Philippines;*
2. *Professor, Institute of Agricultural Engineering, University of the Philippines Los Baños, Laguna 4031, Philippines;*
3. *Professor Institute of Agricultural Engineering, University of the Philippines Los Baños, Laguna 4031, Philippines;*
4. *Research Professor, Postharvest Horticulture Training & Research Center, University of the Philippines Los Baños, Laguna 4031, Philippines*)

Abstract: Near-infrared (NIR) spectroscopy was assessed in predicting maturity and eating quality of 'Carabao' mango fruit. A total of 1,200 fruits were harvested at the green stage at four different harvest dates [100, 110, 120 and 125 days after flower induction (DAFI)]. Fruits were scanned at the green and table-ripe stage (TRS) using NIR reflectance spectroscopy. The fruits were then measured destructively for the determination of dry matter (DM) content at the green stage, total soluble solids (TSS) and sensory attributes at the TRS. The best calibration models were achieved using partial least square regression (PLSR) analysis for predicting DM, TSS and maturity in the short wavelength region of 700-990 nm at 2 nm increment. Principal component analysis-linear discriminant analysis (PCA-LDA) was also used in classifying fruits according to maturity (in terms of DAFI) and eating quality (in terms of overall acceptability or OA). Based on R^2 values, PLSR models are suitable for quality assurance according to maturity ($R^2 = 0.946$, root mean square error of cross validation (RMSECV) = 2.229) and could be used for screening green fruits according to DM ($R^2 = 0.774$, RMSECV = 1.091%). The calibration model for predicting TSS ($R^2 = 0.839$, RMSECV = 1.282) of ripe fruit using NIR spectra at TRS could be used in research but with caution. For classifying fruits according to DAFI and OA, PCA-LDA gave good results using NIR spectra at the green stage with a success rate of 87.8% and 86.0%, 72% and 70% for calibration and validation, respectively. The findings indicate the potential of NIR spectroscopy for non-destructive prediction of maturity and quality parameters of mango. The results of the study could serve as the basis for quality control and automatic sorting system for various commodities.

Keywords: Near-infrared, mango, sweetness, maturity, eating quality, Philippines

Citation: Polinar, Y. Q., K. F. Yaptenco, E. K. Peralta, J. U. Agravante. 2019. Near-infrared spectroscopy for non-destructive prediction of maturity and eating quality of 'Carabao' mango (*Mangifera indica* L.) fruit. Agricultural Engineering International: CIGR Journal, 21(1): 209–219.

1 Introduction

The 'Carabao' mango is one of the leading fruit exports of the Philippines, behind only pineapple and banana. Although production volume has been declining at an average rate of 9,900,000 kg per year since 2002, the value of production has increased annually by Php346 million from 2000 to 2015. Meanwhile, the proportion of fruit exported has decreased from 4.59% in 2002 to 1.75% during the same period (PSA, 2017). Drops in production and quality have been attributed to poor climatic conditions (e.g. typhoons, wind damage) and presence of pests and disease (e.g. leaf hoppers, fruit fly, anthracnose, bacterial wilt). The fresh fruit is exported mainly to the USA, Hong Kong, and Japan (Briones, 2013).

Maturity of 'Carabao' mango at harvest is an

Received date: 2018-05-03 **Accepted date:** 2018-11-01

*** Corresponding author:** Kevin F. Yaptenco, Institute of Agricultural Engineering, University of the Philippines Los Baños, Laguna 4031, Republic of the Philippines. Email: kfyaptenco@up.edu.ph. Tel: +6349-536-3291.

important determinant of postharvest quality. During periods when fruit prices are high, suppliers may harvest early to take advantage of the price situation. For fruit sold to processors, this affects the quality of the raw material and the recovery rate (Digal, 2005). Immature fruit do not ripen evenly and are scored low in terms of sweetness and flavor; furthermore, eating quality is poor, and fruit are susceptible to chilling injury (Agillon, 2003) and development of physiological disorders when exposed to prolonged heat treatment (Esguerra et al., 1990).

Maturity indices at the green stage include counting the number of days after flower induction (DAFI) (110-130 days), flattening of the shoulder and fullness of cheeks, and the proportion of sinkers in 1% salt solution (75% of fruit samples must sink) (Lizada, 1991). Despite its simplicity, the flotation test is often misapplied by mixing an insufficient amount of salt or using only pure water (Esguerra, 2016, personal communication).

Attributes related to eating quality when ripe may include flesh color, firmness, total soluble solids (TSS), pH, dry matter (DM), and titrable acidity (TA). However, the differences between mature and immature fruit based on these attributes is not always clear. DM, specific gravity and sugar content of ‘Mahajanaka’ mango harvested at the green stage were not significantly different at 105-119 days after fruit set (DAFS); at 133-140 DAFS, DM levels were significantly higher and eating quality was considered superior to earlier harvest dates (Saranwong et al., 2004). DM content and flesh color were used as indices of maturity at harvest for Australian varieties of mango (‘KP’, ‘Calypso’, ‘R2E2’, and ‘Celebration’), while TSS correlated with fruit sweetness which is an indicator of the eating quality of mango when ripe (Subedi et al., 2007). However, measurement of these attributes involves laboratory methods which are destructive, time-consuming and laborious.

Near-infrared (NIR) spectroscopy is a rapid and non-destructive method for evaluating internal quality including TSS and firmness in peaches (Kawano et al., 1992), navel orange (Liu et al., 2009), banana (Jaiswal et al., 2012), apple (Ignat et al., 2014) and mango (Schmilovitch et al., 2000; Saranwong et al., 2004; Subedi et al., 2007; Subedi et al., 2013; Jha et al., 2014). The use of NIR technology is expanding due to its cost effectiveness, simplicity of sample preparation, rapid evaluation time and chemical-free measurement. As an alternative to destructive tests for fruit quality, this study examined the use of NIR spectroscopy for determining maturity and quality of fresh ‘Carabao’ mango fruits. The study specifically aimed to (1) characterize physico-chemical properties of ‘Carabao’ mango at different maturities and stages of ripeness, (2) develop NIR calibration models to predict maturity and eating quality of ‘Carabao’ mango, and (3) validate the calibration models.

2 Materials and methods

2.1 Fruit sampling

‘Carabao’ mango fruits were obtained from a single orchard in the town of Sirang Lupa, Calamba in the province of Laguna (Philippines). Fruits were harvested at 100, 110, 120 and 125 DAFI during the peak harvesting season (16 March to 10 April, 2014). At every harvest date, three hundred blemish-free fruits from a set of 15 mango trees (i.e. 20 fruits per tree) were picked manually and packed in plastic crates (20 kg capacity). As an index of maturity, 20 fruits were randomly selected and subjected to flotation test in 1% NaCl solution (Lizada, 1991) to determine the proportion of sinkers (mature) and floaters (immature); a batch with 75% or more of fruit samples classified as sinkers was considered mature. All fruits underwent hot water treatment (HWT) at $52^{o}C$-$55^{o}C$ for 10 minutes, followed by cooling for 10 minutes in tap water (Esguerra and Bautista, 2007). The treated fruits were immediately transported to a laboratory of the Institute of Agricultural Engineering of the University of the Philippines Los Baños. Fruits were allowed to air-dry and equilibrate for one hour at room temperature. Twenty fruits were used for determining fruit density by measuring sample weight and volume using a digital weighing scale and the water displacement method, respectively.

For each harvest date, 300 samples at the green stage of ripeness were scanned by NIR spectrometer and randomly divided into two equal groups (150 samples per group) (Figure 1). In the first group, fruits were subjected to destructive analysis for determining TSS and DM. The second group was ripened at 25°C until the table-ripe stage (TRS) of ripeness was reached. Upon reaching TRS, the fruits were scanned again using NIR spectrometer, followed by analysis for TSS and sensory attributes. Details on the acquisition of NIR spectral data are given below.

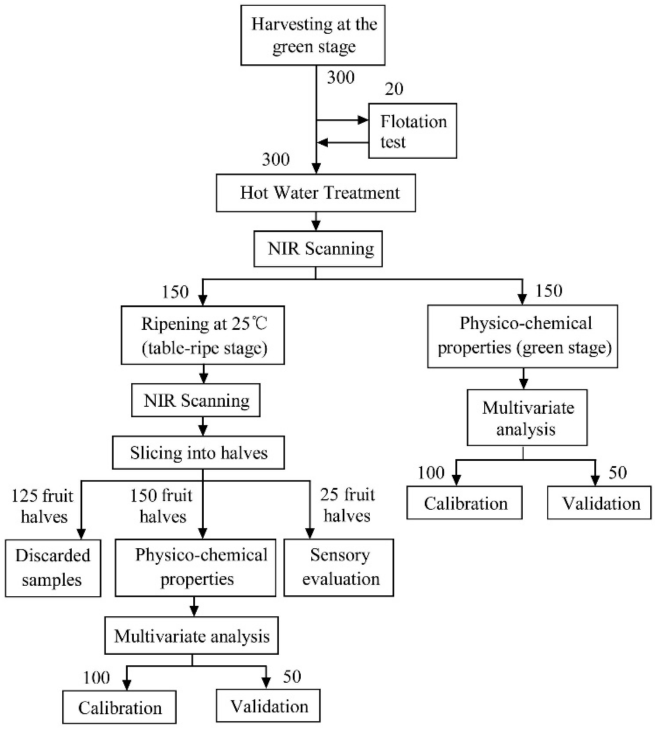

Figure 1 Flowchart of fruit sampling procedure for physico-chemical analysis, NIR spectra acquisition and sensory evaluation of fruit samples for each harvest date. The values represent the number of fruits or fruit halves used

2.2 NIR spectral acquisition

Spectral acquisition was performed using instruments, fiber optics and software manufactured by Ocean Optics Inc (Florida, USA); these included a USB-4000-VIS-NIR spectrometer (operating range of 300-1000 nm) with Spectrasuite software, QR 600-7 VIS125BX fiber optic cables, and a HL-2000 tungsten-halogen lamp (360-2400 nm). The NIR system was allowed to warm up for 60 minutes prior to scanning. Each fruit was placed on a sample holder with a light shield to prevent light interference from the outside environment. A WS1-PTFE reflection standard was used for obtaining a reference spectrum. Samples were scanned 50 times to generate an average spectrum; integration time and boxcar width was set at 50 m sec and 2 nm, respectively. The spectrum was acquired from one cheek of each fruit to produce one sample scan.

2.3 Reference analyses of physico-chemical properties

For determination of physico-chemical properties, the scanned portion of the cheek was excised and chopped into smaller pieces. For TSS determination, a fraction of the chopped sample was crushed to extract juice which was passed through filter paper. TSS was determined from the filtered juice using an Atago RM 45 digital refractometer (0%-45% Brix; Tokyo, Japan) with automatic temperature compensation.

Dry matter content was quantified using the procedure described by Padda et al. (2011). Approximately 5 g of tissue was placed onto filter paper in a Petri dish. The sample was dried at 60°C in a convection oven until a constant weight was reached; drying time was about 2 days. Dry matter content was calculated as the ratio of dry weight to fresh weight, expressed in percent.

2.4 Sensory evaluation

Mango fruits at TRS were evaluated for their eating quality by 25 consumer panelists (Lawless, 1998) of different gender and age. A 10-cm line scale was marked by panelists to evaluate fruits for sweetness and sourness rating. For evaluating the overall acceptability (OA), a nine-point hedonic scale was used ranging from 1 to 9 wherein: 1 = dislike extremely, 2 = dislike very much, 3 = dislike moderately, 4 = dislike slightly, 5 = neither like nor dislike, 6 = like slightly, 7 = like moderately, 8 = like very much, 9 = like extremely. An OA rating of 7 was used as the threshold value in separating acceptable (Pass) and reject (Fail) fruit. Results of sensory analysis were analyzed by ANOVA and Tukey's HSD test at the 5% level of significance using XLSTAT (version 2014.1).

2.5 Statistical analysis

NIR spectral data were transformed to log (1/R), pre-processed [multiple scatter correction (MSC), standard normal variate (SNV)] and analyzed by partial least squares regression (PLSR) with leave-one-out cross-validation. Table 1 gives a summary of the physico-chemical properties and sensory attributes that were analyzed with respect to NIR reflectance data.

The set of spectral data was randomly divided into calibration (n = 100) and validation (n = 50) data sets as described in previous work (Andasuryani et al., 2013; Kawamura et al., 2007; Tsenkova et al., 2001).

Table 1 Summary of PLSR^z analysis performed on physico-chemical properties and sensory attributes versus spectral data

Time of scanning	Stage of Ripeness	
	Full Green	Table-Ripe Stage
At full-green stage	• Days after flower induction • Total soluble solids • Dry matter	• Total soluble solids • Overall acceptability
At table-ripe stage		• Days after flower induction • Total soluble solids • Overall acceptability

Note: ^{z} PLSR – partial least squares regression analysis.

Relative Percentage Deviation (RPD) was determined to assess the robustness or goodness-of-fit of calibration models (Williams and Norris 2001). The guidelines for the interpretation of RPD, R and R^2 are listed in Tables 2 and 3, respectively. Data transformation, pre-processing, PLSR and calculation of RPD values were performed using ParLeS version 3.1 (Viscarra-Rossel, 2008).

Table 2 Guidelines for interpreting relative percentage deviation (RPD)^z

RPD	Classification	Application
0.0-2.3	Very poor	Not recommended for use
2.4-3.0	Poor	Very rough screening
3.1-4.9	Fair	Screening
5.0-6.4	Good	Quality control
6.5-8.0	Very good	Process control
8.1 or higher	Excellent	Any application

Note: ^{z} Source: Williams and Norris (2001).

Table 3 Guidelines for the interpretation of coefficients of correlation (R) and determination (R^2)^z

R	R^2	Interpretation
Up to ±0.5	<0.25	Not usable in calibration
±0.51-0.70	0.26-0.49	Poor correlation, needs further research to identify cause
±0.71-0.80	0.50-0.64	Usable for rough screening
±0.81-0.90	0.66-0.81	Suitable for screening and other approximate calibrations
±0.91-0.95	0.83-0.90	Can be used with caution in most applications, including research
±0.96-0.98	0.92-0.96	Can be used in most applications, including quality assurance
±0.99 or higher	Higher than 0.98	Excellent, can be used in any application

Note: ^{z} Source: Williams and Norris (2001).

2.6 Linear discriminant analysis (LDA)

SAS ver. 9.4 (North Carolina, USA) software was used to perform principal component analysis (PCA) to reduce the variables that are correlated. Retention of principal components was based on a minimum eigenvalue of 1. After PCA, the new set of data comprised of principal components were analyzed using linear discriminant analysis (LDA) to classify fruits according to maturity (i.e. DAFI) and eating quality (i.e. Overall Acceptance) using forward stepwise selection. The significance level for a variable to be retained in the model was set at the default level of 0.15. Table 4 gives a summary of the maturity and sensory attributes that were analyzed by PCA/LDA with respect to NIR reflectance data.

Table 4 Summary of PCA/LDA^z performed on maturity and sensory attributes versus spectral data

Time of scanning	Stage of Ripeness	
	Full Green	Table-Ripe Stage
At full-green stage	• Days after flower induction	• Days after flower induction • Overall acceptability
At table-ripe stage		• Days after flower induction • Overall acceptability

Note: ^{z} PCA – principal component analysis; LDA – linear discriminant analysis.

3 Results and discussion

3.1 Physico-chemical properties of fruit samples

The fruit samples used in the experiment varied in fruit size, volume and density for each harvest date (Table 5). Results showed that fruit size, volume and density tended to increase with fruit maturity. Fruits harvested at 100 and 110 DAFI were considered as immature due to the low ratio of sinkers to floaters. Fruits at 120 and 125 DAFI on the other hand, were considered mature as fruits were 100% sinkers.

Table 5 Effect of harvest date on physical characteristics of 'Carabao' mango fruit

Fruit property^z		Maturity (DAFI)^y			
		100	110	120	125
Weight (g)	Mean	175.0c	229.1b	235.3b	255.4a
	SD	21.1	17.1	18.4	18
	CV	12.00%	7.10%	7.80%	7.10%
Volume (cm^3)	Mean	151.0c	178.5b	189.5ab	203.8a
	SD	21.5	18	22.1	20.5
	CV	14.20%	10.10%	11.60%	10.10%
Fruit density ($g\ cm^{-3}$)	Mean	1.16b	1.29a	1.25a	1.26a
	SD	0.08	0.05	0.06	0.07
	CV	7.10%	5.30%	5.10%	5.30%
Sinker fruit (%)		0c	58.3b	100a	100a

Note: ^{z}Values represent the mean from 80 fruit samples per harvest date. In a row, means with a common letter are not significantly different by Tukey's HSD test at the 5% level of significance. ^{y}DAFI – days after flower induction; SD – standard deviation; CV – coefficient of variation calculated as the ratio of the SD to the mean.

Table 6 shows the physico-chemical properties of fruit samples harvested at different stages of maturity. Fruits harvested at 120 and 125 DAFI had significantly higher amounts of DM content than those fruits harvested at earlier dates. The increase of DM content during maturation is attributed to the continuous accumulation of starch – the main insoluble carbohydrate of DM. When substantial amount of starch is accumulated, a sufficient amount of sugar is hydrolyzed as fruit matures or ripens (Ueda et al., 2000).

Table 6 Effect of harvest date on physico-chemical properties[z] of 'Carabao' mango fruit[y]

Maturity (DAFI)	TSS (green)	DM (green)	TSS (ripe)	Sensory attributes		
	(°Brix)	(%)	(°Brix)	Sweetness	Sourness	Overall acceptability
100	5.7±0.4c	12.5±1.2d	9.1±1.1d	2.74b	2.78b	5.25c
110	5.8±0.4c	15.0±1.5c	11.9±1.4c	4.11b	3.67ab	6.30b
120	6.0±0.4b	16.7±1.3b	14.9±1.4b	6.99a	2.74b	8.25a
125	6.9±0.7a	17.7±1.5a	16.8±1.7a	6.16a	4.88a	7.60a
Range	4.6-8.3	6.6-23.0	7.2-20.9	-	-	-

Note: [z] DAFI – days after flower induction; TSS – total soluble solids; DM – dry matter. [y] Values represent the mean±SD from 150 fruit samples per harvest date. In a column, means with a common letter are not significantly different by Tukey's HSD test at the 5% level of significance.

The eating quality improved as fruit matured up to 120 DAFI; no significant differences were observed between 120 and 125 DAFI for OA and sweetness. At 125 DAFI, the significant increase in sourness may be due to increase titrable acidity during maturation (Tirtosoekotjo, 1985). Based on the threshold value of OA = 7, fruits at 100 and 110 DAFI were classified as "Fail" while fruits at 120 and 125 DAFI were classified as "Pass".

3.2 Near-infrared spectra analysis

The shape of spectral curves differed between green and fully-ripe fruits in the 550-700 nm range (Figure 2 and 3). The absorbance at 675 nm was dominant in green fruits (shown as a valley in the reflectance plot) that may correspond to their chlorophyll content at 680 nm. A relatively flat zone was observed for all spectra above 750 nm, which is common to plants (Williams and Norris, 2001). A peak can be observed at around 970 nm that is related to water and sugar O-H stretching (second overtone of OH stretch at 960, third combination of overtone at 840 nm) and C-H stretching (third overtone stretch at 910 nm) (Golic et al., 2003). Figure 2 and 3 show a comparison of the original reflectance spectra of fruit samples (green and ripe) and spectra treated by multiplicative scatter correction (MSC), respectively.

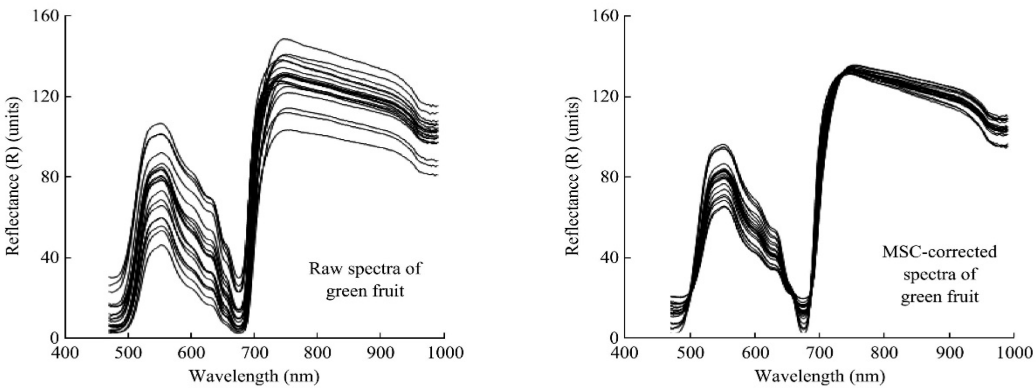

Figure 2 Original spectra (left) and corrected spectra (right) of green 'Carabao' mango fruit. Corrected spectra were obtained by multiplicative scatter correction (MSC)

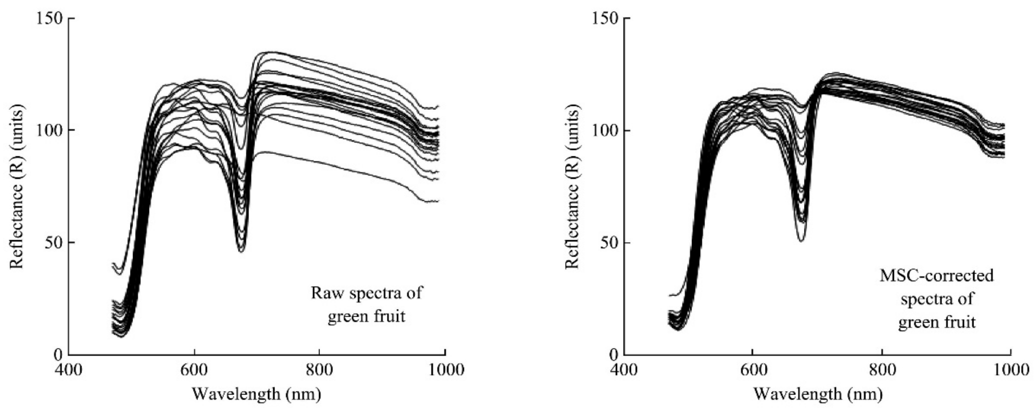

Figure 3 Original spectra (left) and corrected spectra (right) of ripe 'Carabao' mango fruit. Corrected spectra were obtained by multiplicative scatter correction (MSC)

3.3 Prediction of green-stage fruit quality and maturity using green-stage NIR spectra

After NIR scanning of harvested green fruits, thin slices of flesh (~1 mm thickness) were removed to determine the dry matter content (Padda et al. 2011). It resulted in an acceptable PLSR calibration model with the highest R^2 value of 0.774 and RMSECV of 1.091% (Table 7) in the NIR wavelength range 700-990 nm. The wavelength range corresponds to higher values of R^2 and RPD, and a lower RMSECV without overfitting. The best calibration model was obtained without any pre – processing technique. Based on its R^2 value, the model is suitable for screening and other approximate calibrations but the RPD (2.10) implies that the model is very rough for screening. Scatter plots of green-stage DM predicted using green-stage spectra are shown in Figure 4; the results could be due to the small range of DM at harvest such that any differences between maturities were difficult to detect by NIRS. Furthermore, for this study, the peel was removed after NIR scanning (i.e. peel was not included in the DM determination) which could have resulted in low R^2 of the calibration model. Other causes could be moisture loss after peeling and non-uniform thickness of the excised peel layer (Saranwong et al., 2004).

Table 7 Performance of partial least squares regression models^z at different wavelength ranges for prediction of green-stage dry matter ‘Carabao’ mango fruit using green-stage spectra

Wavelength range (nm)	Treatment	Number of factors	Calibration (Leave-one-out)			Validation		
			R^2	RMSECV	RPD	R^2	RMSECV	RPD
470-990 (VIS-NIR)	No correction	15	0.732	1.190	1.93	0.734	1.162	1.90
	MSC	15	0.728	1.202	1.91	0.725	1.180	1.87
	SNV	15	0.731	1.195	1.92	0.730	1.169	1.89
470-700 (VIS)	No correction	10	0.633	1.391	1.65	0.628	1.348	1.64
	MSC	10	0.633	1.394	1.65	0.639	1.333	1.66
	SNV	10	0.632	1.396	1.64	0.642	1.324	1.67
700-990 (NIR)	No correction	9	0.774	1.091	2.10	0.728	1.198	1.84
	MSC	8	0.752	1.143	2.01	0.709	1.238	1.78
	SNV	8	0.706	1.259	1.82	0.657	1.395	1.58

Note: ^{z} RMSECV – root mean square error of cross-validation; RPD – relative percentage deviation; VIS – visible; NIR – near-infrared; MSC – multiplicative scatter correction; SNV – standard normal variate.

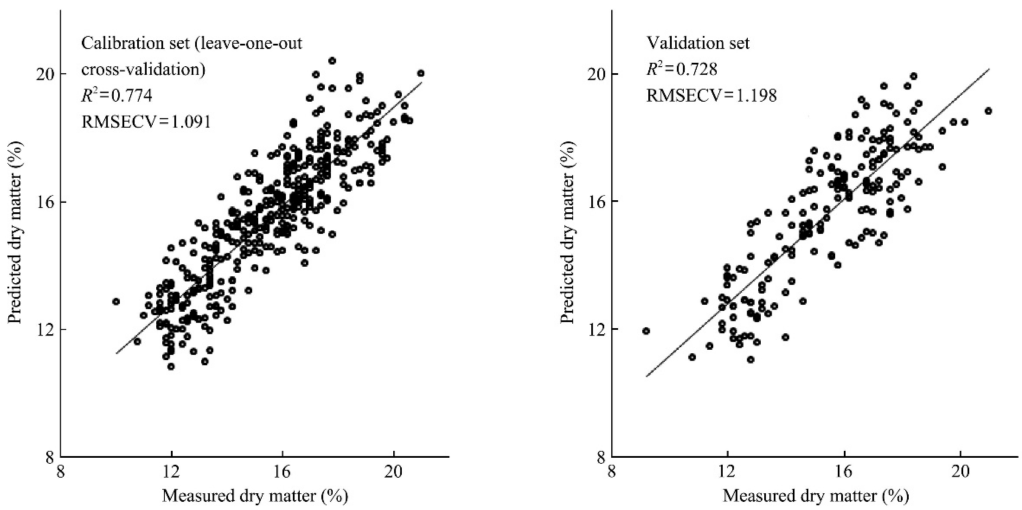

Figure 4 Performance of PLSR model for predicting dry matter (DM) in the wavelength range of 700-990 nm for calibration (left) and validation (right) sets while green using the spectra at the green stage

The TSS model did not result in an acceptable calibration at green stage (R^2=0.448; RMSECV=0.454%; RPD = 1.34, data not shown). Similar results for TSS at harvest (R^2 = 0.62) for different varieties of mango (‘Kensington Pride’, ‘Calypso’, ‘R2E2’ and ‘Celebration’) from Australia were obtained by Subedi et al. (2007).

For the prediction of maturity based on DAFI, calibration model (R^2 = 0.946; root mean square error of cross validation (RMSECV) = 2.229 days; RPD = 4.32) applied with multiple scatter correction (MSC) provided a better result (Table 8) that could be used for research and screening of fruit. For example, fruit could be classified into immature and mature fruit using a threshold value based on the percentage of floaters (e.g. DAFI when the proportion of floaters <25%). However, some overlap between predicted harvest dates can be observed (Figure 5) which could result in misclassified fruit.

Table 8 Performance of PLSR models at different wavelength ranges for prediction of maturity (in terms of DAFI) of green 'Carabao' mango fruit using green-stage spectra

Wavelength range (nm)	Treatment	Number of factors	Calibration (Leave-one-out)			Validation		
			R^2	RMSECV	RPD	R^2	RMSECV	RPD
470-990 (VIS-NIR)	No correction	15	0.896	3.111	3.09	0.909	2.899	3.32
	MSC	15	0.922	2.681	3.59	0.929	2.585	3.73
	SNV	15	0.924	2.649	3.63	0.931	2.533	3.80
470-700 (VIS)	No correction	10	0.751	4.809	2.00	0.726	5.070	1.90
	MSC	10	0.813	4.159	2.31	0.825	4.058	2.37
	SNV	10	0.814	4.145	2.32	0.826	4.012	2.40
700-990 (NIR)	No correction	10	0.942	2.317	4.15	0.952	2.125	4.52
	MSC	10	0.946	2.229	4.32	0.947	2.225	4.32
	SNV	10	0.816	4.428	2.17	0.833	4.273	2.25

Note: ^{z} RMSE – root mean square error of cross-validation; RPD – relative percentage deviation; VIS – visible; NIR – near-infrared; MSC – multiplicative scatter correction; SNV – standard normal variate.

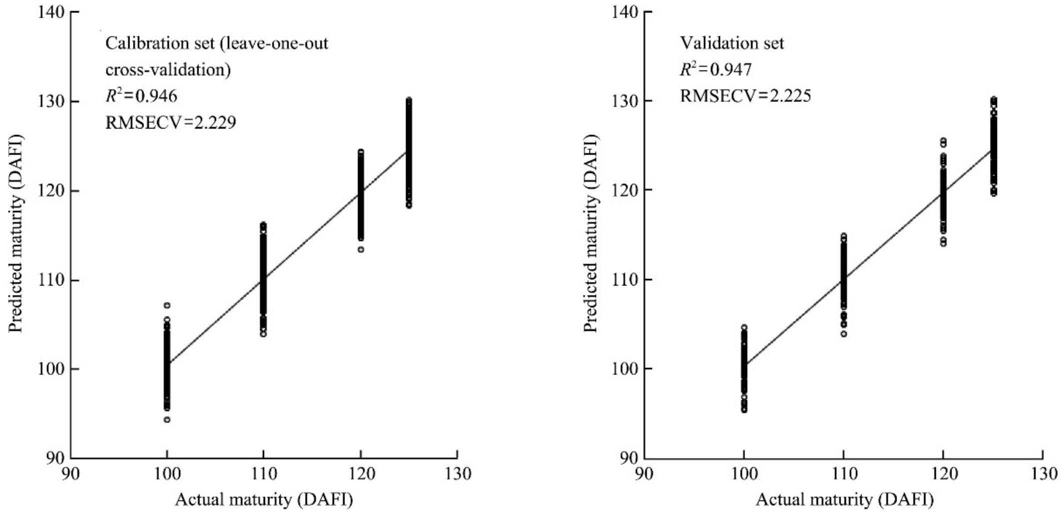

Figure 5 Performance of PLSR model MSC-corrected for predicting maturity (DAFI) in the wavelength range of 700-990 nm for calibration (left) and validation (right) sets while green using the spectra at the green stage

3.4 Prediction of ripe-stage TSS and maturity using ripe-stage NIR spectra

At the market end, consumers often purchase mango fruit based on appearance, only to find out that the eating quality is inferior. This is common for immature fruits that have been ripened artificially; even though such fruit have developed the full yellow peel with uniform color distribution, the taste is bland. A technique for rapidly classifying fruit into process grades and premium grades would ensure customer satisfaction, encourage repeat sales, and discourage early harvesting.

Saranwong et al. (2004) analyzed NIR spectra using PLSR at the 700-1100 nm range to predict eating quality of 'Mahajanaka' mango at the green stage of ripeness. A calibration model using both dry matter and starch was sufficiently accurate to predict TSS (which is a measure of eating quality). Similarly, Subedi and Walsh (2007) used PLSR analysis of NIR spectra at the 300-1150 nm range for predicting eating quality of Australian mangoes at harvest. For the latter study, the use of starch content for predicting eating quality was removed since DM and starch were correlated with each other. This simplified the calibration stage since the reference method for starch content was difficult. The results were considered sufficiently accurate to predict eating quality at the ripe stage from green fruit, with a validation R^2 of 0.92). Neto et al. (2017) used NIR spectra (306-1140 nm) to predict TSS of 'Palmer' mango; the best prediction was made using PLSR over the 699-999 nm range (R^2 = 0.87, RMSECV=1.39%).

For the present study, the PLSR calibration model for TSS at TRS using spectra at TRS was reasonably accurate (R^2=0.839; RMSECV=1.282%; RPD=2.53, data not shown). The calibration model was improved after treating spectral data with the standard normal variate (SNV) technique in the VIS-NIR wavelength range of 470-990 nm (Table 9). Since the typical consumer is capable of detecting differences in sweetness of approximately 1% Brix (Saranwong et al. 2004), some improvement in the model is needed to be able to achieve better prediction of TSS. A scatterplot of predicted versus measured TSS shown in Figure 6) demonstrates the performance of the PLSR model for the calibration and prediction sets. The range of TSS for ripe 'Carabao' fruit samples was 15%-21% Brix. Yaptenco et al. (2013) reported a TSS range of 9.2%-15.2% Brix for 'Carabao' mango at the table ripe stage, harvested at 100-125 DAFI; the best eating quality was at 120 DAFI with TSS of 15% Brix. Based on R^2 and RPD value, the calibration model of the present study could be used as a rough classifier with 15% Brix as a threshold value to sort fruit into two grades based on sweetness.

Table 9 Performance of PLSR^z models at different wavelength ranges for prediction of total soluble solids of table-ripe 'Carabao' mango fruit using ripe-stage spectra

Wavelength range (nm)	Treatment	Number of factors	Calibration (Leave-one-out)			Validation		
			R^2	RMSECV	RPD	R^2	RMSECV	RPD
470-990 (VIS-NIR)	No correction	15	0.826	1.334	2.39	0.844	1.302	2.42
	MSC	15	0.839	1.283	2.49	0.852	1.269	2.48
	SNV	15	0.839	1.282	2.49	0.853	1.244	2.53
470-700 (VIS)	No correction	10	0.683	1.799	1.78	0.723	1.742	1.81
	MSC	10	0.716	1.701	1.88	0.736	1.686	1.87
	SNV	10	0.720	1.689	1.89	0.744	1.643	1.92
700-990 (NIR)	No correction	10	0.824	1.342	2.38	0.804	1.407	2.24
	MSC	8	0.833	1.305	2.45	0.813	1.380	2.28
	SNV	8	0.833	1.306	2.45	0.804	1.430	2.20

Note: ^{z} PLSR – partial least squares regression; RMSE – root mean square error of cross-validation; RPD – relative percentage deviation; VIS – visible; NIR – near-infrared; MSC – multiplicative scatter correction; SNV – standard normal variate.

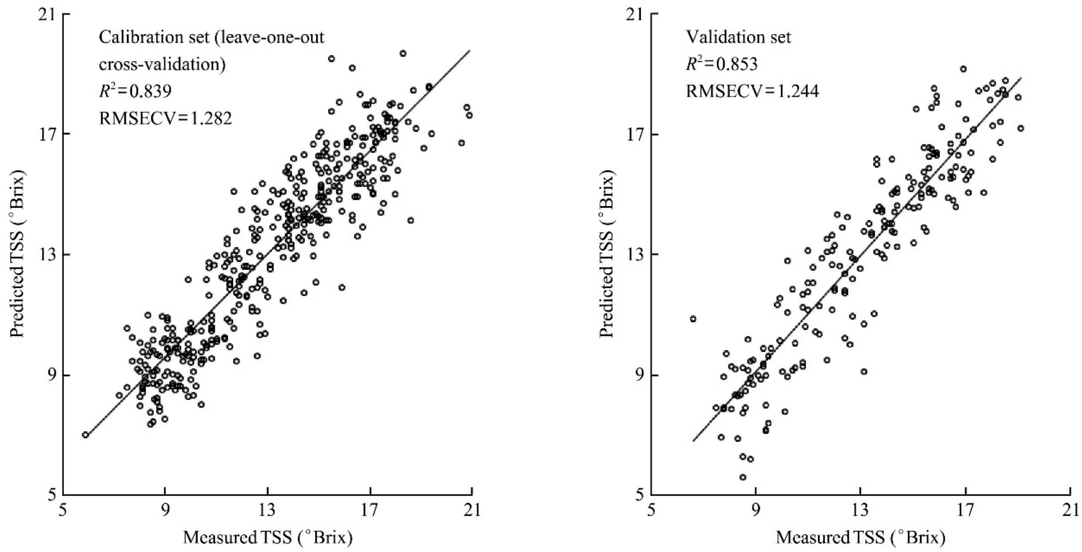

Figure 6 Performance of PLSR model MSC-corrected for predicting total soluble solids (TSS) in the wavelength range of 470-990 nm for calibration (left) and validation (right) sets when ripe using the spectra at the ripe stage

3.5 Linear discriminant analysis

Linear discriminant analysis also provided good results in classifying fruit at the green stage according to maturity (DAFI). For this data set, four principal components with eigenvalues greater than 1 were retained. Eigenvector values showed a similar grouping of scores at 914-916 nm (corresponding to dry matter), 962-964 nm (water) and 978-980 nm (sugars) (Williams and Norris 2001). The values in the diagonal of the classification table showed 87.8% and 86.0% of the fruits correctly classified for calibration and validation sets, respectively (Table 10). The separations according to maturity can be clearly observed in Figure 7, which plots the values of the three discriminant functions (dimensions) for each harvest date. From the data plot, it seems that 125 and 100 DAFI are relatively separated while 110 and 120 DAFI are less separable.

Acceptable results were also obtained for classifying green fruit based on eating quality with a threshold value of OA = 7; 72% and 70% of samples were correctly classified for calibration and validation sets, respectively (Table 11). Four principal components were also retained to separate fruit according to OA (Figure 8). Only a single discriminant function was needed to separate fruit according to overall acceptability.

Linear discriminant analysis was less successful at classifying ripe fruit according to maturity and eating quality. While both parameters were successfully predicted at 70% and 81% of samples, respectively, for the calibration set, the success rate for validation was only 31% and 52%, respectively.

Table 10 Summary of results for classifying green-stage 'Carabao' mango according to maturity by linear discriminant analysis

Actual DAFI	Predicted DAFI for calibration			
	100	110	120	125
100	88.5	0.0	11.5	0.0
110	0.0	82.3	15.2	2.5
120	0.5	13.5	86.0	0.0
125	0.0	6.0	0.0	94.0
Total correctly classified = 87.8%				
Actual DAFI	Predicted DAFI for validation			
	100	110	120	125
100	88.9	1.0	10.1	0.0
110	0.0	83.0	16.0	1.0
120	0.0	23.0	77.0	0.0
125	0.0	5.0	0.0	95.0
Total correctly classified = 86.0%				

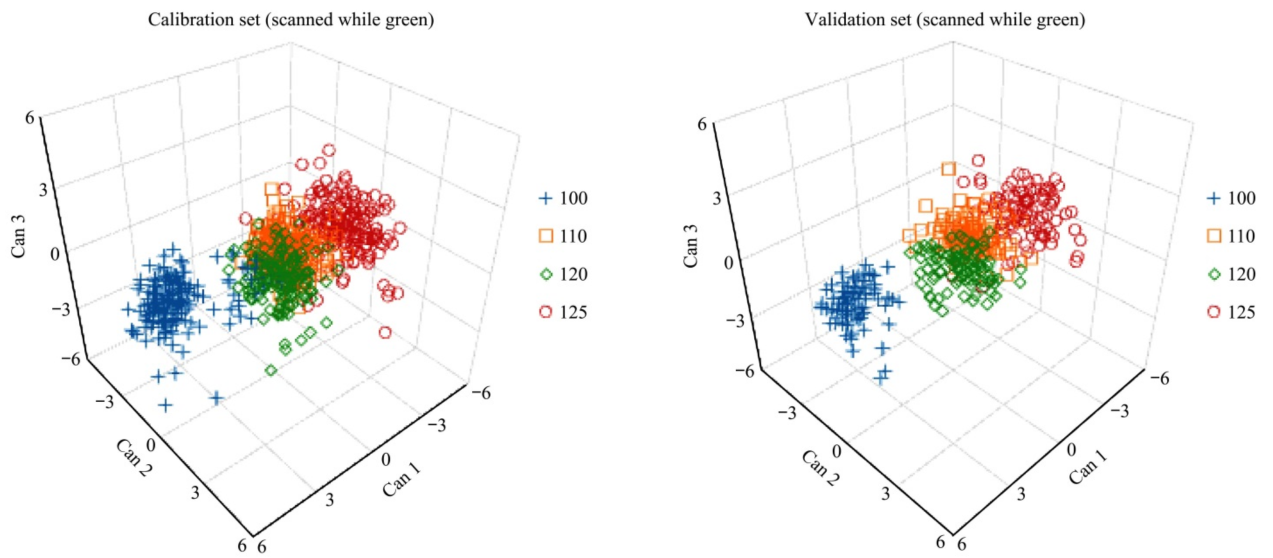

Figure 7 Performance of discriminant analysis for classifying green 'Carabao' mango according to maturity (in terms of DAFI) for calibration (left) and validation (right) set of samples

Table 11 Summary of results for classifying green-stage 'Carabao' mango according to eating quality by linear discriminant analysis

Actual OA[z]	Predicted OA for calibration		Predicted OA for validation	
	Fail	Pass	Fail	Pass
Fail	65.6%	34.4%	61.8%	38.2%
Pass	20.8%	79.2%	22.0%	78.0%
Total correctly classified[y]	72.4%		69.9%	

Note: [z] OA – overall acceptability; fruit samples with OA > 7 were rated as "Pass". [y]Computed as the sum of correctly classified samples (in %) rated as Fail and Pass divided by 200, expressed in %.

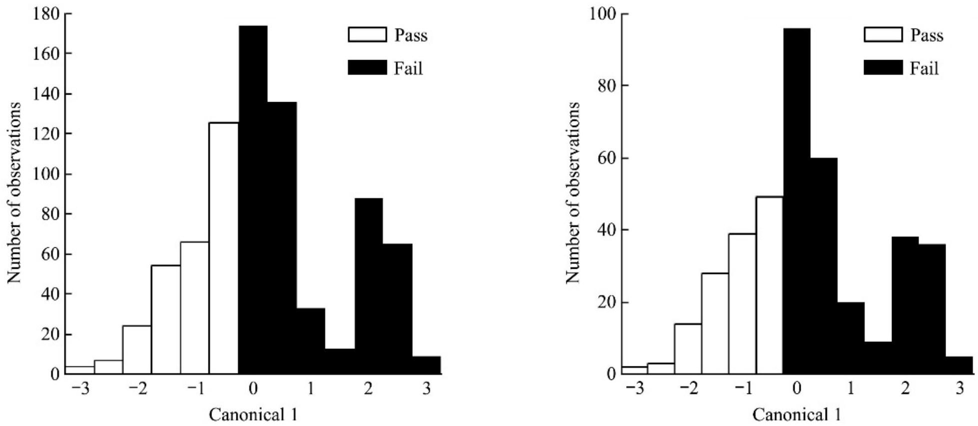

Figure 8 Performance of discriminant analysis for classifying green 'Carabao' mango according to predicted eating quality (in terms of OA) for calibration (left) and validation (right) sets

4 Summary and conclusion

Based on the R^2 values, the partial least squares regression models for predicting dry matter (DM) (R^2=0.774), TSS (R^2=0.839) and maturity (R^2=0.946) were suitable for quality control while the RPD value for the prediction of DM (RPD=2.10) suggest that model may be rough for screening. Linear discriminant analysis (LDA) provided good results in classifying fruits according to maturity (DAFI) and eating quality (overall acceptability or OA). For classifying fruits according to DAFI using NIR spectra at the green stage, success rate showed 87.8% and 86.0% of the fruits correctly classified for calibration and validation sets, respectively. Acceptable results were also obtained for classifying green fruit based on eating quality with a threshold value of OA = 7 which reflects 72% and 70% of the fruits correctly classified for calibration and validation sets, respectively.

The findings indicate the potential of near infrared (NIR) spectroscopy for non-destructive prediction of maturity and quality parameters of mango. The results of the study would allow the growers to optimize their income by harvesting fruits only at the best time when fruits have reached the proper maturity. Most importantly, this could serve as the basis for pricing and quality control and automatic sorting system for various commodities. Other potential applications proposed by Subedi et al. (2013). for this method include on-tree and on-farm tracking and mapping of fruit maturation, checking and training of fruit harvesters, and as a tester to target selected markets.

Acknowledgement

This study was supported by the Philippine Department of Science and Technology – Engineering Research & Development for Technology (DOST-ERDT) scholarship program and is based on the Masteral thesis of the first author. The technical assistance provided by Dr. Paul R. Armstrong and Engr. Elizabeth B. Maghirang of the US Department of Agriculture – Stored Product Insect and Engineering Research Unit on data analysis is gratefully acknowledged; financial support for travel and training was provided by the DOST-Philippine Council for Industry, Energy and Emerging Technology R&D (PCIEERD) through its short-term training program "Bridging the Human Resource Competency Gaps in Support of the National R&D Agenda". Instruments for NIR spectroscopy were obtained through a research grant provided by the Philippine Commission on Higher Education (CHED) entitled "Machine Vision and Infrared Spectroscopy Applications in Agricultural Engineering."

References

Agillon, A. B. 2003. Physiological and biochemical responses of `Carabao' mango (*Mangifera indica* L.) fruit to chilling stress. M.S. thesis. University of the Philippines: Los Baños, Laguna, Philippines.

Andasuryani, A., Y. A. Purwanto, I. W. Budiastra, and K. Syamsu. 2013. Determination of catechin content in gambir powder from dried gambir leaves quickly using FT NIR PLS model. *International Journal on Advanced Science, Engineering and Information Technology*, 4(5): 1–5.

Briones, R. M. 2013. Market Structure and Distribution of Benefits From Agricultural Exports: The Case of The Philippine Mango Industry. PIDS Discussion Paper Series No. 2013–16.

Available at: http://hdl.handle.net/10419/126937. Accessed 6 July 2017.

Digal, L. N. 2005. Benefit Diffusion and Linkage Development in the Philippine Tropical Fruits Sector. Makati City, Philippines: Asian Institute of Management. World Bank / National Economic Development Authority. Available at: http://siteresources.worldbank.org/INTPHILIPPINES/Resources/Digal-word.pdf. Accessed 4 May 2018.

Esguerra, E. B., and O. K. Bautista. 2007. Disease Control. In *Postharvest Technology for Southeast Asian Perishable Crops*, eds. E. B. Esguerra, and O. K. Bautista, 337–363. Philippines: University of the Philippines Los Baños.

Esguerra, E. B., S. R. Brena, M. U. Reyes, and M. C. C. Lizada. 1990. Physiological breakdown in vapour heat-treated 'Carabao' mango. In *Symposium on Tropical Fruit in International Trade 269*, 425–434. Honolulu, Hawaii, USA: International Society of Horticultural Science.

Golic, M., K. B. Walsh, and P. Lawson. 2003. Short-wavelength near-infrared spectra of sucrose, glucose, and fructose with respect to sugar concentration and temperature. *Applied Spectroscopy*, 57(2): 139–145.

Ignat, S., J. Lurie, J. Nyasordzi, V. Ostrovsky, H. Egozi, A. Hoffman, H. Friedman, A. Weksler, and Z. Schmilovitch. 2014. Forecast of apple internal quality indices at harvest and during storage by VIS/NIR spectroscopy. In *Food and Bioprocess Technology*. New York: Springer Science Business Media.

Jaiswal, P., S. N. Jha, and R. Bharadwaj. 2012. Non-destructive prediction of quality of intact banana using spectroscopy. *Scientia Horticulturae*, 135: 14–22.

Jha, S. N., P. Jaiswal, K. Narsaiah, M. Gupta, R. Bharadwaj, R. Kumar, and R. Sharma. 2014. Non-destructive prediction of maturity of mango using near infrared spectroscopy. *Journal of Food Engineering*, 124: 152–157.

Kawamura, S., M. Kawasaki, H. Nakatsuji, and M. Natsuga. 2007. Near infrared spectroscopic sensing system for online monitoring of milk quality during milking. *Sensing and Instrumentation for Food Quality and Safety*, 1(1): 37–43.

Kawano, S., H. Watanabe, and M. Iwamoto. 1992. Determination of sugar content in intact peaches by near infrared spectroscopy with fiber optics in interactance mode. *Journal of the Japanese Society for Horticultural Science*, 61(2): 445–451.

Lawless, H., and H. Heymann. 1998. *Sensory Evaluation of Food: Principles and Practices*. New York, USA: Chapman and Hall.

Liu, Y., X. Sun, and H. Zhang. 2009. Nondestructive assessment of quality of nanfeng mandarin fruit by a portable near infrared spectroscopy. *International Journal of Agricultural and Biological Engineering*, 2(1): 65–71.

Lizada, M. C. C. 1991. Postharvest Physiology of Mango: A Review. *Acta Horticulturae*, 291: 437–453.

Neto, J.O.S., M.W.D. de Assis, I.P. Casagrande, L.C.C. Junior, and G.H.A. Teixeira. 2017. Determination of 'Palmer' mango maturity indices using portable near infrared (VIS-NIR) spectrometer. Postharvest Biology and Technology. 130(): 75–80.

Padda, S., C. Amarante, R. Garcia, D. Slaughter, and E. Mitcham. 2011. Methods to analyse physico-chemical changes during mango ripening: A multivariate approach. *Postharvest Biology and Technology*, 62(3): 267–274.

PSA [Philippine Statistics Authority]. 2017. CountrySTAT

Philippines. Available at: http://countrystat.psa.gov.ph/. Accessed 5 July 2017.

Saranwong, S., J. Sornsrivichai, and S. Kawano. 2004. Prediction of ripe-stage eating quality of mango fruit from its harvest quality measured nondestructively by near infrared spectroscopy. *Postharvest Biology and Technology*, 31(2): 137–145.

Schmilovitch, Z., A. Mizrach, A. Hoffman, H. Egozi, and Y. Fuchs. 2000. Determination of mango physiological indices by near-infrared spectrometry. *Postharvest Biology and Technology*, 19(3): 245–252.

Subedi, P. P., K. B. Walsh, G. Owens. 2007. Prediction of mango eating quality at harvest using short-wave near infrared spectrometry. *Postharvest Biology and Technology*, 43(3): 326–334.

Subedi, P., K. Walsh, and P. Purdy. 2013. Determination of optimum maturity stages of mango using fruit spectral signatures. *Acta Horticulturae*, 992: 521–527.

Tirtosoekotjo, R. A. 1985. Ripening Behavior and Physico-Chemical Characteristics of 'Carabao' Mango (*Mangifera indica* L.) treated with Acetylene from Calcium Carbide. Ph.D. diss., University of the Philippines: Los Baños, Laguna, Philippines.

Tsenkova, R., S. Atanassova, S. Kawano, and K. Toyoda. 2001. Somatic cell count determination in cow's milk by near-infrared spectroscopy: A new diagnostic tool. *Journal of Animal Science*, 79(1): 2550–2557.

Viscarra-Rossel, R. A. 2008. ParLeS: Software for chemometric analysis of spectroscopic data. *Chemometrics and Intelligent Laboratory Systems*, 90(1): 72–83.

Williams, P. C., and K. Norris. 2001. Implementation of near-infrared technology. In *Near-Infrared Technology in Agricultural and Food Industries*, eds. P. Williams, and K. Norris, 145–169. St. Paul, Minnesota: American Association of Cereal Chemists.

Ueda, M., K. Sasaki, N. Utsonomiya, K. Inaba, and Y. Shimabayashi. 2000. Changes in physical and chemical properties during maturation of mango fruit (*Mangifera indica* L. 'Irwin') cultured in a plastic greenhouse. *Food Science and Technology Research*, 6(4): 299–305.

Yaptenco, K. F., M. P. Angeles, E. B. Esguerra, E. P. Serrano, M. A. Cho and S. T. Choi. 2013. Nondestructive firmness measurement as a maturity index for 'Carabao' mango (*Mangifera indica* L.) fruit. *Philippine Agriculturist*, 96(1): 55–65.