Predictions of fruit shelf life and quality after ripening: Are quality traits measured at harvest reliable indicators?

Thibault Nordey[1,2], Fabrice Davrieux[3,5], Mathieu Léchaudel[4,5]

[1]TheWorld Vegetable Center, Eastern and Southern Africa, P.O. Box 10, Duluti, Arusha, Tanzania

[2]CIRAD, UPR Hortsys, 34398 Montpellier, France

[3]CIRAD, UMR QUALISUD, F-97455 Saint Pierre, La Réunion, France

[4] CIRAD, UMR QUALISUD, F-97130 Capesterre-Belle-Eau, Guadeloupe, France

[5]QualiSud, Univ Montpellier, CIRAD, Montpellier SupAgro, Univ d'Avignon, Univ de La Réunion, Montpellier, France.

Abstract

Nondestructive methods such as near infrared spectroscopy (NIRS) are increasingly used in sorting lines to assess quality traits of unripe fruit, i.e. dry matter (DM) and total soluble solid (TSS) contents, in order to create homogenous batches of fruit. The use of this approach is based on the assumption that fruit quality traits at harvest are reliable indicators of their post-harvest behavior and their quality after ripening. The present study tested this assumption by analyzing the relationships between quality traits at harvest and after ripening. In parallel, models were developed to determine the capacity of NIRS measurements on unripe fruit at harvest to predict their shelf life and quality after ripening.

The quality traits DM, TSS content, pulp color (PC) and titratable acidity (TA) of 92 mangoes from different harvests, production years, and orchards were compared at harvest and after ripening. Previously developed NIRS models were used to nondestructively assess the quality traits of the mangoes at harvest. New partial least squares (PLS) regressions using different variable selection procedures and preprocessing techniques were used to predict fruit shelf life and fruit quality after ripening based on NIRS measurements at harvest.

Weak relationships ($r^2 < 0.41$) were found between fruit quality traits measured at harvest and after ripening, except for DM content ($r^2 = 0.61$). The PC of mango measured at harvest was found to be the best indicator of fruit shelf life. Errors of PLS regressions to predict the TSS content (RMSEV = 1.1%), titratable acidity (RMSEV = 0.52%), and the Hue angle of the flesh (RMSEV = 1.86 °) were in the same range as those of linear regressions based on quality traits assessed at harvest except for PC. This work provides evidence that fruit maturity and quality should be assessed using different indicators.

Keywords: *Mangifera indica*; Near-infrared spectroscopy; Non-destructive prediction; Eating quality; PLSR

Introduction

The heterogeneity of quality and maturity of fruit at harvest is a widespread problem in numerous species that needs to be addressed all along the supply chain to reduce postharvest losses and to insure constant quality for consumers. After harvest, fruit are generally sorted and graded to create homogenous batches based on the assumption that their post-harvest behavior and quality will be similar after ripening. To improve quality assessments based on visual rating, i.e. absence of defects, size, color and shape, several nondestructive methods have been developed to assess other fruit quality descriptors at the time of measurement such as total soluble solid (TSS) content, titratable acidity (TA), and the dry matter (DM) content. These methods include the electronic nose (Lebrun *et al.*, 2008), near infrared spectroscopy (NIRS) (Jha *et al.*, 2012a; Nordey *et al.*, 2017; Saranwong *et al.*, 2004; Subedi *et al.*, 2007), visual spectroscopy (Jha *et al.*, 2006), and specific gravity (Kapse and Katrodia, 1996). Sorting fruit using quality traits measured at harvest assumes that the fruit composition at this stage is a reliable descriptor of its quality after ripening, and of its shelf life, i.e. the length of the period between harvest and the ripe fruit stage. This assumption relies on the fact that the quality of ripe fruit is determined at harvest since the accumulation of dry matter and water in fruit stops once the fruit is picked. Fruit dry matter contains the preliminary metabolites and precursors of secondary metabolites that undergo considerable changes during fruit ripening and hence determine the quality of ripe fruit. In a few days, ripening processes increase fruit quality to its optimum, which then decline until the fruit become inedible due to over ripening. Metabolic pathways of preliminary and secondary metabolites are controlled by a balance of different phytohormones, including ethylene, abscisic acid and gibberellins. The metabolism of these phytohormones and their involvement in ripening processes are used to differentiate climacteric fruit from non-climacteric fruit. Managing fruit shelf life is essential to insure optimum fruit quality for consumers, especially in the case of highly perishable climacteric fruit such as mango, banana and avocado. The shelf life and the quality of fruit after ripening are known to be closely related to their stage of maturity at harvest since the shelf life of fruit harvested early is longer but their quality is

65 reduced, i.e. they are smaller, have a lower sugar content, paler pulp, than fruit harvested later (Joas
66 *et al.*, 2012; Nordey *et al.*, 2016).
67 Although the quality of fruit varies considerably with their stage of maturity, this does not mean that
68 fruit composition is a reliable indicator of fruit maturity since the concentration of primary and
69 secondary metabolites is known to vary considerably depending on the growing conditions, e.g.
70 irrigation, the fruit to leaf ratio, and the position of the fruit in the canopy (Léchaudel and Joas,
71 2007). For this reason, several studies on mango (Lechaudel *et al.*, 2010), papaya (Urbano Bron *et al.*,
72 2004), and apple (Song *et al.*, 1997) preferred to use the optical proprieties of chlorophyll in the fruit
73 peel assessed with a fluorometer as an indicator of fruit maturity rather than fruit quality descriptors.
74 Although several studies on mango (Saranwong et al., 2004; Subedi et al., 2007), apple (Palmer *et al.*,
75 2010), and kiwifruit (Jordan *et al.*, 2000; McGlone *et al.*, 2002b) focused on the relationship between
76 DM content and TSS content at harvest and after ripening, few investigated relationships with other
77 quality traits such as TA and pulp color (PC), which are also of importance in consumers' perception
78 of quality.
79 The first objective of the present study was to investigate the validity of the assumption that fruit
80 quality descriptors measured at harvest are reliable indicators of the shelf life of fruit and of their
81 quality after ripening. Mango was used as a model since numerous studies have underlined the
82 capacity of NIRS measurements to non-destructively measure several fruit quality traits in mango:
83 TSS content, dry matter content, titratable acidity and pulp color (Cortés *et al.*, 2016; Jha *et al.*,
84 2012b; Marques *et al.*, 2016; Nagle *et al.*, 2010; Nordey et al., 2017; Rungpichayapichet *et al.*, 2016;
85 Schmilovitch *et al.*, 2000). We took advantage of previously developed NIRS models to analyze the
86 relationships between the quality traits measured at harvest and after ripening in a set of mango
87 fruit sampled from different orchards, harvests and production years. Although numerous studies on
88 mango focused on the use of NIRS to measure fruit quality traits (see above mentioned studies) and
89 maturity (Cortés et al., 2016; Nagle et al., 2010; Rungpichayapichet et al., 2016; Subedi et al., 2007)
90 at the time of measurement, only a few investigated the potential of NIRS measurements at harvest

91 to predict the quality of ripe fruit (Subedi et al., 2007) and shelf life. The second aim of this study was
92 thus to evaluate the accuracy of NIRS measurements for such predictions. The results of this study
93 should help stakeholders of fruit value chains choose reliable indicators to assess fruit shelf life and
94 quality after ripening.
95

Material and methods

Samples

A total of 92 mango fruit (*Mangifera indica* cv. 'Cogshall') harvested during the 2010–2011 and 2014–2015 production seasons in four orchards in the northwest, west and southwest of Reunion Island (20◦52'48''S, 55°31'48''E) were used. Tree size, spacing and ages differed between orchards, as did fertilization, irrigation and pruning. Fruit were harvested between 90 and 120 days after full bloom to account for the wide range of variation in the stage of maturity at harvest from the green mature to the yellow point stage considered as the onset of fruit ripening for Cogshall mangoes (Lechaudel et al., 2010). A NIR spectrum was collected for each fruit at harvest, after which the fruit was weighed and left to ripen at 20 °C and 80% relative humidity (RH).

The mangoes were destroyed for composition analysis after ripening. To ensure that ripe fruit was the same physiological age for analysis, respiratory metabolism and climacteric rise were used as indicators. Previous studies on the Cogshall cultivar (Joas *et al.*, 2009; Joas et al., 2012; Joas *et al.*, 2010) showed that the fruit quality traits TSS content and TA, firmness vary according to the climacteric stage of the fruit. In line with these studies, mangoes were considered to be ripe with correct quality and taste three days after they had reached their highest respiration rate. Respiration rates were measured daily on each fruit by placing the mango in an individual 3 L airtight jar, and CO_2 concentration was measured at 20 min intervals for 1 hour by gas chromatography using an Agilent M200 instrument (SRA, Marcy l'Etoile, France).

Measurements of fruit quality

At the ready to eat stage, mango cheeks were cut off longitudinally to measure the PC with a Minolta Chroma Meter CR300 (Konica Minolta, Osaka, Japan) and described using the Hue angle criterion. Variations in TA, DM content and TSS content within mangoes (Nordey *et al.*, 2014) were taken into account through measurements made on a puree obtained by blending the fruit flesh in a Grindomix blender (Retsch, Haan, Germany). Fresh juice was extracted by filtering the puree through gauze to

measure the TSS content using an ATC-1E refractometer (Atago, Tokyo, Japan) and TA. TA, expressed as mass percentage of citric acid (%), was measured using an automated titrimeter (TitroLine easy, Schott, Mainz, Germany) with a 0.05 mol L^{-1} NaOH solution. The DM content of the flesh was calculated from the dry mass measured after lyophilization compared with fresh mass.

Chemometrics

At harvest, NIR spectra measurements were collected on the surface of the fruit near the apex over the 600–2300 nm wavelength range using a portable spectrometer equipped with a contact probe (LABSPEC 2500, Analytical Spectral Devices, Inc., Boulder, CO, USA). In line with our previous studies NIR measurements were made on the fruit apex since peel color changes in this part of the fruit is used as an indicator of the fruit maturity for cogshall mangoes (Lechaudel et al., 2010; Nordey et al., 2017).

NIR measurements were used to non-destructively measure the fruit quality traits DM, TA, TSS and PC at harvest using previously developed partial least square (PLS) models (Nordey et al., 2017). The accuracy of the models was expected to be satisfactory since they were calibrated on mangoes taken from similar orchards in the same year of production as the ones used in the present study. Spectral measurements collected at harvest were also used to predict the shelf life of the fruit and their quality after ripening by establishing new PLS models.

Samples were divided into calibration and validation sets at an 80:20 ratio for each quality trait evaluated, i.e., DM and TSS content, PC and TA, by random sampling on percentiles of the quality attribute values. Partial least squares regressions (PLSR) were established using the PLS package (Mevik and Wehrens, 2007) of the R software (R Development Core Team, 2012) using the methodology described by (Cornillon, 2010). The number of PLSR factors was determined to reduce the prediction error by cross validation on 20% of the calibration set using the mean square error of prediction as an indicator.

Several spectral data pre-processing and variable selection methods developed in our previous study (Nordey et al., 2017) were tested to improve the prediction performance of PLSR. The preprocessing

147 methods tested were first and second derivatives using the Savitzky-Golay smoothing filter with a
148 second-order polynomial and a 10-nm window size using the prospectr package (Stevens and
149 Ramirez-Lopez, 2013). Interval partial least square (IPLS) regressions, associated with the stepwise
150 and the backward methods, were performed to select the combination of wavelength windows that
151 best predicted performance. Algorithms for IPLS regressions were designed following the
152 methodology presented by Andersen and Bro (2010). As proposed by Nicolaï *et al.* (2007), the root
153 mean square error (RMSE) was used as an indicator to evaluate the predictive performance of PLSR
154 using the calibration (RMSEC) and prediction (RMSEP) datasets.

155 Statistical analysis

156 Covariance analyses were performed to assess the impacts of growing conditions, i.e. years of
157 production and orchards, on relationships between quality traits measured at harvest and after
158 ripening.
159 A principal component analysis (PCA) was performed to analyze variations in raw NIR spectra
160 collected on fruit at harvest using the FactoMineR package (Lê *et al.*, 2008). Fruit shelf life was
161 plotted as a supplementary categorical variable and the positions of the shelf life categories were
162 plotted on the PCA plot with their confidence ellipses at 95%. Simple and multiple linear regressions
163 were calibrated and tested using the same calibration and prediction data sets as those used for
164 PLSR. A variable selection procedure was applied to the accuracy of multiple linear regressions
165 following the methodology suggested by Cornillon (2010) based on the LEAPS package (Lumley and
166 Miller, 2009) and on the Bayesian information criterion (BIC). The root mean square error (RMSE)
167 was used as an indicator to evaluate the predictive performance of linear regressions for the
168 calibration (RMSEC) and prediction (RMSEP) datasets. A relative RMSEP was calculated as the ratio
169 between the RMSEP and the mean of all measurements.
170

Results

Changes in fruit quality attributes between harvest and after ripening

Figure 1 shows the relationships between the fruit quality traits TSS content, dry matter content, TA and PC measured at harvest using NIRS spectra and after ripening using destructive measurements. Results revealed marked variations in quality at harvest since fruit weight varied between 170 g and 665 g (data not shown), TSS content varied between 4.5 and 20%, TA varied between 2.25 and 12.22%, DM content varied between 12.2 and 23.9%, the hue angle of the PC varied between 83.5 and 116.7 °. Fruit were ripe from two to 17 days after harvest and their fresh mass varied between 156 and 637 g (data not shown), DM content varied between 10.8 and 21%, TA varied between 0.35 and 4.35%, TSS content varied between 10.2 and 22%, and the hue angle of the PC varied between 80.15 and 92.7 °. Weak relationships were found ($r^2 < 0.41$) between quality traits at harvest and after ripening, except for DM content ($r^2 = 0.61$). The TSS content in ripe fruit was correlated ($r^2 = 0.67$, Figure 1E) with the DM content measured at harvest, in contrast to TA (Fig. 1F) and PC (Figure 1G).

The fruit shelf life was related to the PC ($r^2 = 0.7$, Figure 1J) and to the TSS content ($r^2 = 0.62$, Fig. 1H) and to a lesser extent to TA ($r^2 = 0.5$, Figure 1I) and to the DM content ($r^2 = 0.45$, Figure 1K) measured at harvest. All relationships established between quality traits measured at the harvest and after ripening were found to vary significantly with growing conditions, i.e. the year of production and/or the orchard.

The accuracy of linear regressions between quality traits at harvest and after ripening is shown in Table 1. The variable selection procedure made it possible to increase the accuracy of multiple linear regressions to predict fruit quality traits after ripening. This approach showed that PC and DM content of the ripe fruit were best predicted using DM content at harvest as the only indicator. Although the TSS content in ripe fruit was well predicted using DM content at harvest as the only explanatory variable, our results showed that including TA and PC in the multiple linear regression slightly increased prediction accuracy. TA of fruit after ripening was found to be best predicted using

PC and TA measured at harvest. In line with previous results, PC at the harvest was shown to be the best indicator of fruit shelf life.

Use of NIRS measurements at harvest to predict the quality of ripe fruit and shelf life

NIRS spectra measured on fruit from 600 nm to 2,300 nm at harvest (Figure 2A) were used to predict their shelf life at 20 °C and 80% RH, as well as their quality traits after ripening.

Reflectance spectra acquired at harvest varied with the shelf life of the fruit (Figure 2A) and a principal component analysis on raw NIR spectra was performed to highlight these variations (Figure 2B). Principal component analysis revealed more variation in the NIR spectra acquired on fruit with a longer shelf life.

PLSR were developed to predict the quality of fruit after ripening and their shelf life at harvest using NIR measurements. The results of the preliminary analyses displayed in Figure 3A to 3E underline the difference in the capacity of NIR windows to predict fruit quality traits and shelf life. These figures also show that quality traits in ripe fruit are linked to different regions in the NIR spectra. Different data preprocessing methods (first and second derivative) as well as variable selection procedures (IPLS backward and stepwise) with different sized windows in the NIR spectra (10, 25, 50 and 100) were used to increase the prediction accuracy of PLSR (Table 2). The models with the least prediction errors were selected for the calibration and validation datasets. Models with similar accuracy but fewer factors were selected to increase the robustness of the results. In line with Figure 3A to 3E, different regions in the NIR spectra were selected in the models to predict quality traits (Figure 3F to 3J). Predictions of the TSS content in ripe fruit were found to rely on reflectance measurements at harvest from 1,000 nm to 1,200 nm, as well as on reflectance measurements around 1,800 nm. Similar results were found for DM content since reflectance measurements around 1,000 nm were selected by the variable selection procedure to predict this trait. Predictions of PC and TA in ripe fruit were both related to measurements in the NIR region from 1,600 to 1,800 nm. Reflectance

measurements in the visible region (around 800 nm) were found to be of importance only for the prediction of fruit shelf life.

Prediction accuracies of the selected PLSR are shown in Figure 4. A RMSEP of 1.1%, 0.52%, 1.86 °, 1.26% and 1.78 days were found for the TSS content, TA, the hue angle of the PC, DM content and the shelf life, respectively. Marked discrepancies were found between the accuracy of models since relative RMSEP of 6.9%, 46%, 2.1%, 8%, 18.3%, and 18.3% were obtained for the TSS content, TA, the hue angle of the PC, DM content, and the shelf life, respectively.

Errors of the same order of magnitude were obtained when predicting quality traits and shelf life using PLSR and linear regressions based on the quality traits assessed at harvest, except for the PC, i.e. RMSEV = 1.86 ° versus 3.17 °.

Discussion

Are fruit quality traits at harvest reliable indicators of fruit shelf life and quality after ripening?

Our results show that the color of the pulp is a good indicator of fruit shelf life (Fig. 1J, Table 1). This result is in line with previous measurements made on mango by Subedi *et al.* (2007), who reported that fruit maturity was better correlated with PC ($r^2 = 0.79$) than with DM content ($r^2 = 0.66$). Previous studies showed that the color of mango flesh is closely linked with its carotenoid contents (Vasquez-Caicedo *et al.*, 2005), mostly represented by all-trans-carotene, all-trans-violaxanthin, and 9-cis-violaxanthin (Litz, 2009; Rosalie *et al.*, 2015). The biosynthetic carotenoid pathway is known to be triggered during fruit ripening leading to marked changes in the color of the mango flesh. The better capacity of PC to predict the fruit shelf life than the other traits studied could be explained by the lower sensitivity of the carotenoid metabolism to fruit growing conditions than the sensitivity of sugars and acids (Joas et al., 2012; Rosalie et al., 2015) and by the impact of phytohormones in the carotenoid metabolism that drive fruit ripening (McAtee *et al.*, 2013).

Not surprisingly, our results showed that fruit DM contents at harvest and after ripening were closely correlated. Although the composition of fruit DM undergoes major changes during ripening, its content varies only slightly due to water losses and fruit respiration (Nordey *et al*., 2016). During ripening, the starch that accumulates in mangoes throughout their development on the tree is converted into soluble sugars, i.e., saccharose, glucose and fructose, thereby increasing the fruit TSS content (Léchaudel *et al.*, 2005). In line with the results of previous studies (Saranwong *et al*., 2004; Subedi *et al., 2007)*, our results indicated that TSS content at the harvest is not a reliable indicator of the TSS content in fruit after ripening, which is better predicted by DM content at harvest. Several modeling approaches have been developed on mango (Léchaudel *et al.*, 2007), peach (Lescourret *et al.*, 2011), and tomato (Liu *et al.*, 2007) to predict changes in fruit DM during fruit growth and ripening. These models predict the DM composition of fruit by simulating changes in the fruit maturity stage and its dry mass balance. Empirical relationships used in the modeling approaches developed on mango can roughly predict mango glucose, fructose and sucrose contents,

and malic, citric, pyruvic and oxalic acid contents, since correlation coefficients (r^2) obtained between predictions and observations ranged between 0.43 and 0.66 (Léchaudel *et al.*, 2007). One of the main problems involved in predicting changes in the composition of fruit DM is simulating the impacts of ripening. Further work combining modelling approaches to simulate the metabolism of phytohormones involved in fruit ripening (Génard and Gouble, 2005) and their impacts on metabolic pathways of primary and secondary metabolites is thus needed to better predict changes in the DM composition of the fruit during ripening.

In contrast to TSS content, fruit TA after ripening was poorly correlated with predicted fruit DM content at harvest (Figure 3F). Numerous organic acids are responsible for variations in TA in mango, but citric and malic acids are known to have the most influence (Léchaudel et al., 2005; Medlicott and Thompson, 1985). Some modelling approaches have also been developed to simulate TA and the pH in fruit flesh during fruit growth and ripening (Etienne *et al.*, 2013; Lobit *et al.*, 2003). These approaches are hampered by the number of organic acids in fruit and by the lack of knowledge on the mechanisms involved in their metabolism and storage. These models succeeded in underlining the close relationship between organic acid metabolism and fruit respiration. This relationship was used by our team to hypothesize that the observed variations in TA among mangoes after ripening can be partly explained by differences in the climacteric respiratory crisis observed between fruit, depending on their stage of maturity at harvest (Nordey et al., 2016). Interestingly, the multiple linear relationships we established in the present study (Table 1) reinforce this hypothesis, since, as mentioned above, TA in the fruit after ripening was better predicted using both TA and PC at harvest, and the latter was the best indicator of fruit maturity (Table 1, Figure 1J).

Like TA, PC after ripening was poorly correlated with DM content and PC at harvest. This is in agreement with the results obtained by Joas *et al.* (2012), who already underlined the lack of proportionality between the carotenoid content in fruit at harvest and in ripe fruit (Figure 1C). In contrast to DM and TSS contents, these authors reported that the carotenoid content in mango flesh at harvest did not vary either with the fruit carbon supply (Joas et al., 2012) or with the fruit water

supply (Rosalie et al., 2015) but did vary with the stage of maturity at harvest (Joas et al., 2012). The impact of carbohydrate availability in fruit on the metabolism of carotenoids was discussed by Poiroux-Gonord *et al.* (2012), who suggested that carotenoid biosynthesis was not promoted by higher concentrations of carbohydrate precursors. Our results confirm their hypothesis, since PC after ripening was not correlated with TSS or DM content at harvest.

Finally, the results of the present work confirm that fruit DM content at harvest is a reliable indicator of TSS content in ripe fruit, which is known to be closely correlated with their sugar content. Nondestructive measurements such as specific gravity and NIRS have already been successfully used to accurately predict the DM content of several fruit species including mango (Nordey et al., 2017; Saranwong et al., 2004) and kiwi (Jordan et al., 2000; McGlone et al., 2002b). In the present study, all relationships between fruit quality traits at harvest and after ripening were found to vary with fruit growing conditions, i.e. with the orchard and/or year of production. To avoid the need to develop specific relationships for each growing condition, the robustness of these linear regressions could be could be improved by including samples of several seasons and growing regions within the calibration.

Our results also showed that DM content at harvest was not a reliable indicator of TA or PC after ripening. The PC at harvest was found to be the best indicator of fruit shelf life. TA in ripe fruit was found to be linked to PC and TA at harvest, suggesting that it varied with the stage of maturity at harvest. Our results underline the fact that although the stage of maturity of fruit and their quality are closely related, they should not be assessed using the same indicators.

Use of NIRS to predict fruit quality after ripening and shelf life at harvest

Unlike other nondestructive measurements such as weight or density, NIRS spectra are collected in specific locations in the fruit. Like in previous studies (Lechaudel et al., 2010; Nordey et al., 2017) NIR measurements were made on the fruit apex, whereas in other studies, measurements were made on the mango shoulders (Saranwong et al., 2004), or in the center of the fruit cheek (Rungpichayapichet

et al., 2016), or at several different locations (Jha *et al.*, 2014; Marques et al., 2016). Since marked variations in both mango quality and maturity were measured previously (Nordey et al., 2014), we would have expected predictions of quality and maturity to vary according to the position the measurements were made on the fruit. In contrast to previous studies (Nordey et al., 2014; Saranwong et al., 2004), NIR measurements collected in the present study were used to predict the quality and maturity of the fruit as a whole and not of the fleshy part of the measuring area. It is so assumed through the approach used in the present study that quality and maturity in the apex part of the fruit are reliable indicators of the quality and maturity of the whole mango. It is worth noting that automation of the proposed method would be hampered by the need of make NIR measurements at a specific position on the fruit. However, this challenge could be overcome by developing new models based on several NIR spectra randomly collected on the fruit surface.

The accuracy of predictions of fruit quality after ripening made at harvest using NIRS spectra (Table 2) was found to be of the same order of magnitude as linear regressions based on the prediction of quality attributes at harvest, except for the color of the pulp, i.e. RMSEV = 1.86 ° versus 3.17 °.

In contrast to other quality attributes, the accuracy of PLSR to predict TSS content in fruit after ripening was lower than the accuracy of PLSR previously developed to predict the fruit quality at the time of measurement: 1.1% versus 0.6%. This can be explained by the smaller difference in quality attributes between ripe fruit than between unripe fruit harvested at different stages (from green mature to fully ripe).

Like in other fruit, mango spectra were dominated by a water spectrum with overtone bands of OH bonds at 970, 1450 nm and a combination band at 1940 nm (Figure 2)(Nicolaï et al., 2007). The near infrared spectrum of mango is also composed of overtones and combination bands of organic compounds. In line with previous studies, NIR measurements made at harvest at around 1000 nm played an important role in predicting dry matter content and TSS content in ripe mangoes. This region of the NIR spectra was linked to overtone starch at 990 nm. This result supports the results

previously obtained by Saranwong *et al.* (2004) suggesting that the starch content of mango at harvest is a good indicator of TSS content in ripe fruit.

PLSR using NIR spectra at harvest predicted DM content (Fig. 4D) and TSS content (Figure 4A) in ripe fruit better than PC (Figure 4C) and TA (Figure 4B). Our results confirm the conclusions of previous studies concerning the limited accuracy of NIR models to predict TA in mangoes that may be hampered by the number of different organic acids in this species as well as by changes in the ratio of the two main organic acids during ripening (Marques et al., 2016; Nordey et al., 2017; Schmilovitch et al., 2000). Similar results have also been reported in apple (McGlone *et al.*, 2002a) and in passion fruit (Maniwara *et al.*, 2014). The TA and PC of ripe fruit were found to be best predicted in PLSR using NIR measurements at 1600-1800 nm. Previous studies using NIRS showed that the β carotene content in mango (Rungpichayapichet *et al.*, 2015) and Chinese kale (Chen *et al.*, 2009) was related to absorbance of around 1750 nm. This is in agreement with linear regressions showing that titratable acidity in ripe fruit is linked to PC at harvest.

The NIR models developed in the present study succeeded in predicting fruit shelf life with an average error of less than two days. These results are satisfactory compared with the measurement error of shelf life using fruit respiration, which is around one day. The results in Figure 4J show that the region of the spectrum near 800 nm is important to predict fruit shelf life. This region is related to absorption by chlorophyll pigments, which are known to be a reliable descriptor of mango maturity (Lechaudel et al., 2010). The chlorophyll content in mango peel is known to increase during the first stages of mango development and to decrease during fruit ripening (Medlicott *et al.*, 1986). Although several authors used NIR measurements to predict the stage of maturity of mangoes (Cortés et al., 2016; Nagle et al., 2010; Rungpichayapichet et al., 2016; Subedi et al., 2007), to our knowledge, this is the first report on the use of NIRS to predict fruit shelf life at harvest. It should be noted that the fruit shelf life of fruit predicted in the present study is for storage at 20 °C and 80% RH. In any other post-harvest conditions, PLSR would need to be recalibrated to predict fruit shelf life. Our results provide evidence that NIR models can help predict some quality traits of ripe fruit,

i.e. dry matter, TSS content and shelf life. Future studies should use more samples to improve the robustness and the accuracy of the models, especially for predictions of TA of ripe fruit at harvest.

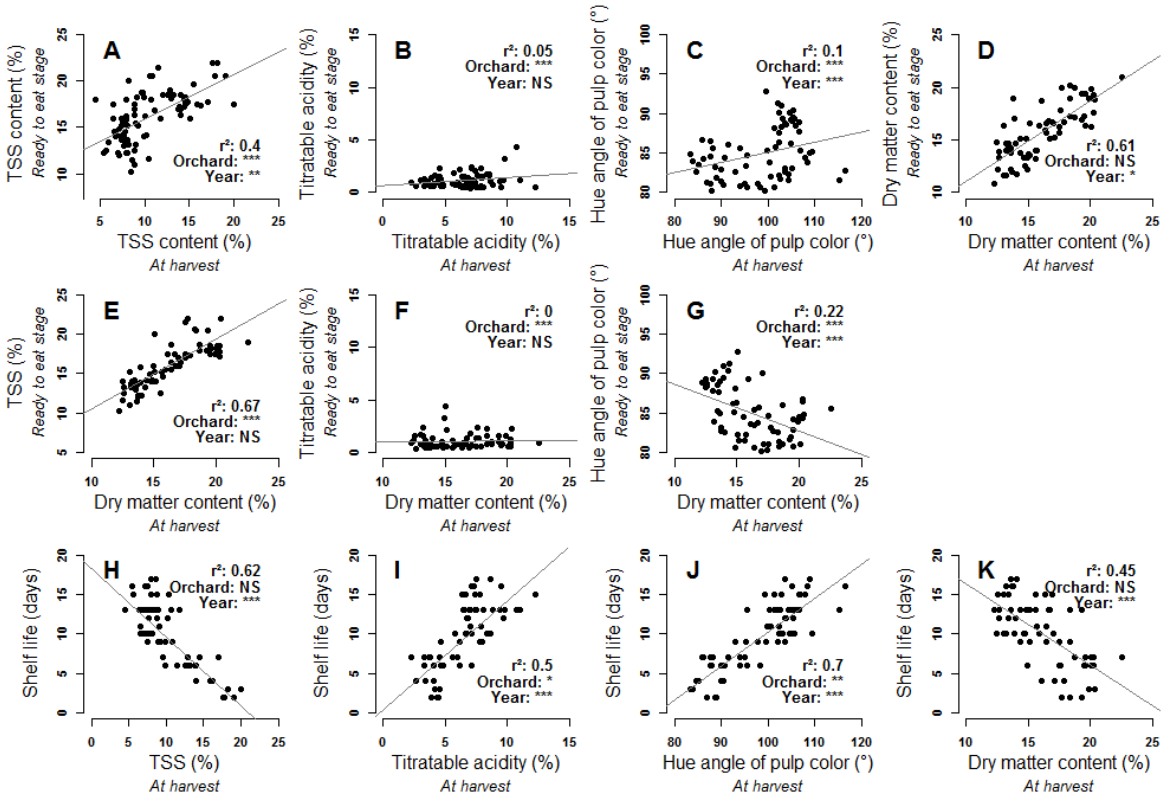
Conclusions

The quality and the maturity of fruit are two notions that are often confused since similar indicators are used to assess them. The present work used NIR models to analyze the relationship between mango quality traits at harvest and after ripening. Our results provide evidence that fruit DM content at harvest is a useful indicator of TSS content in fruit after ripening but not of TA or PC. Pulp color at harvest was found to be the best indicator of fruit shelf life because of its relative insensitivity to growing conditions. The NIR models we developed enabled prediction of fruit shelf life, TSS content and DM content in ripe fruit. Prediction accuracy was nevertheless lower for fruit acidity and PC.

386

387

388 **Tables and figures**



389

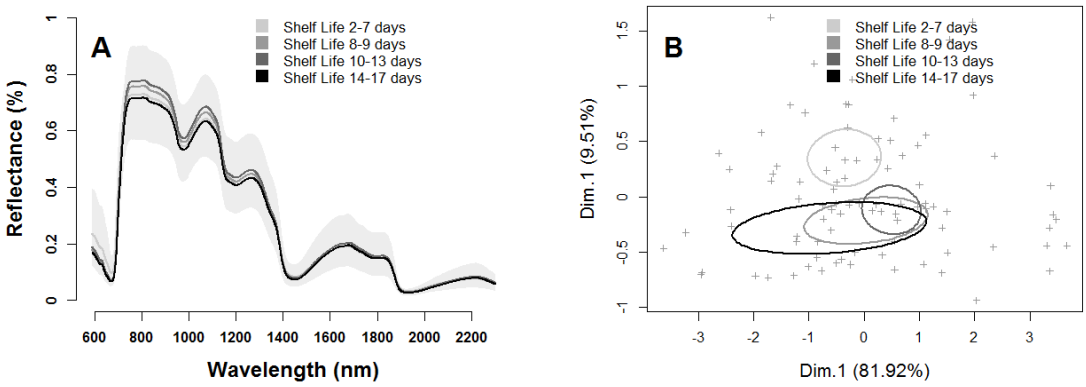
390 **Figure 1**: Relationship between (i) quality traits measured at harvest and after ripening (A to D), (ii)

391 quality traits measured after ripening and dry matter content measured at harvest (E to G) and (iii)

392 shelf life and quality traits measured at harvest (H to K). Correlation coefficients (r^2) are indicated in

393 the figures and asterisks indicate whether years and orchard significantly impact relationships

394 displayed with p. values of 0.001 ***, 0.01**, and 0.05*.



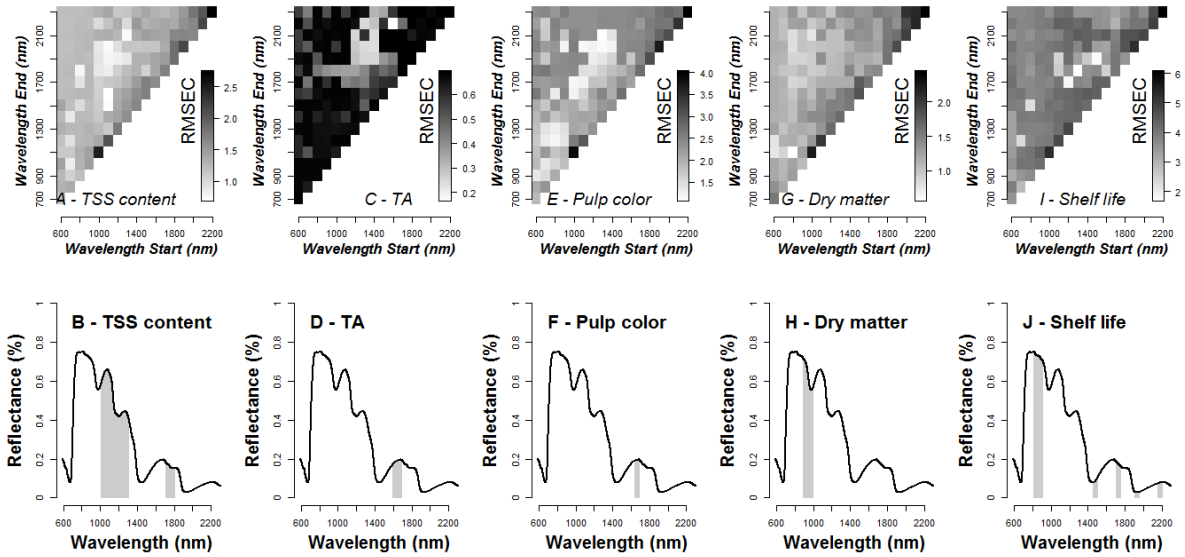
395

396 **Figure 2**: Raw NIR spectra acquired on the peel at the apex of the mango fruit at harvest with

397 averages calculated by range of self-life (A) and their scores in the principal component analysis (B).

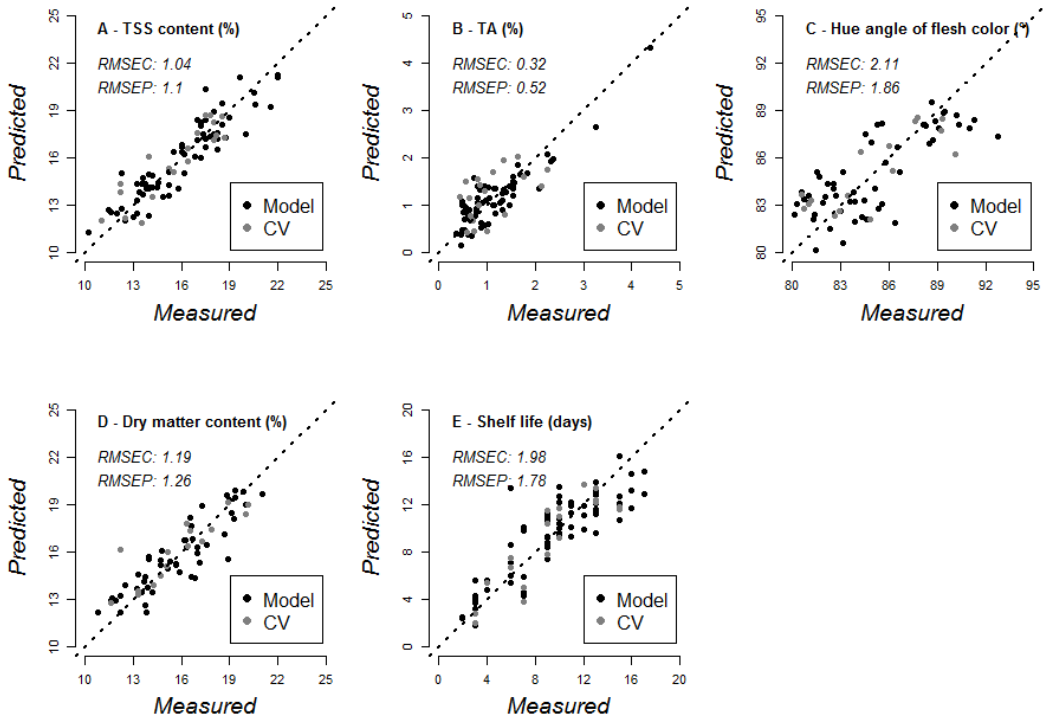
398 Average fruit shelf life is shown as supplementary categorical variables with their confidence ellipses

399 at 95%.



400

401 **Figure 3:** Prediction performances of the selected PLSR regressions, in terms of root mean square
402 standard error of calibration (RMSEC), for different wavelength windows (from a starting point to an
403 end point) and the NIR regions selected for percentage total soluble solid (TSS) content (A-B),
404 percentage titratable acidity (TA) (C-D), percentage pulp color (PC) (E-F), percentage dry matter
405 (DM) content (G-H) and shelf life (in days) in the best models (I-J).



406

407 **Figure 4**: Accuracy of selected partial least square (PLS) regressions in predicting total soluble solid

408 (TSS) content (A), titratable acidity (TA) (B), pulp color (C), dry matter content (D) and shelf life (E)

409 with calibration and validation data sets.

410

411 **Table 1**: Accuracy of linear regressions based on quality attributes measured at harvest: percentage
412 titratable acidity (TA), percentage dry matter content (DM), percentage TSS content, and hue angle
413 of the pulp color (in °) to predict the shelf life and quality of fruit after ripening.

Quality traits in ripe fruit	Quality traits measured at harvest	RMSEC	RMSEP
Titratable acidity	*Titratable acidity*	0.13	0.09
	Dry matter content	0.13	0.1
	TSS content	0.13	0.1
	Pulp color	0.13	0.1
	All	0.1	0.16
	Pulp color & titratable acidity	**0.12**	**0.07**
Pulp color	*Titratable acidity*	3.21	3.60
	Dry matter content	**2.84**	**3.17**
	TSS content	3.14	3.30
	Pulp color	3.13	3.23
	All	2.16	3.91
Dry matter content	*Titratable acidity*	2.26	1.84
	Dry matter content	**1.67**	**1.06**
	TSS content	2.03	1.78
	Pulp color	1.84	1.50
	All	1.22	1.62
TSS content	*Titratable acidity*	2.47	2.07
	Dry matter content	1.53	1.22
	TSS content	2.13	1.82
	Pulp color	1.91	1.49
	All	1.06	1.57
	Titratable acidity & Dry matter content & Pulp color	**1.41**	**1.18**
Shelf life	*Titratable acidity*	2.83	2.57
	Dry matter content	3.04	2.08
	TSS content	2.55	2.05
	Pulp color	**2.27**	**1.56**
	All	1.91	1.79

414

415

416

Table 2: Capacity of partial least squares regressions (PLSR) to predict quality of fruit after ripening and their shelf life at harvest using NIR spectra with different variable selection and preprocessing methods. The root mean square error (RMSE) was used as an indicator to evaluate the predictive performance of PLSR for calibration (RMSEC) and prediction (RMSEP) datasets.

		DM (%)			TSS content (%)			Titratable acidity (%)			Hue angle of pulp color (°)			Shelf life (days)		
		RMSEC	RMSEV	Factors	RMSEC	RMSEV	Factors	RMSEC	RMSEV	Factors	RMSEC	RMSEV	Factors	RMSEC	RMSEV	Factors
Raw spectra	*No variable selection*	0.91	1.57	9.00	1.12	1.42	11.00	0.17	0.68	20.00	2.14	2.90	5.00	2.02	1.75	12.00
	IPLS_Backward_10	1.45	1.77	8.00	1.93	1.68	3.00	0.66	0.50	1.00	2.74	2.76	3.00	3.43	2.62	3.00
	IPLS_Backward_25	1.44	1.42	4.00	2.14	2.21	3.00	0.67	0.52	1.00	2.78	3.12	3.00	3.36	2.29	2.00
	IPLS_Backward_50	1.40	2.05	7.00	2.19	2.13	5.00	0.67	0.51	1.00	1.72	2.75	7.00	3.26	2.61	4.00
	IPLS_Backward_100	**1.19**	**1.26**	**5.00**	1.54	1.44	6.00	0.46	0.46	6.00	2.21	2.57	4.00	2.78	1.94	5.00
	IPLS_Stepwise_10	1.41	1.46	7.00	1.11	1.69	10.00	0.52	0.66	7.00	1.99	3.93	7.00	2.50	1.48	7.00
	IPLS_Stepwise_25	1.25	1.83	6.00	1.09	2.01	11.00	0.22	0.54	15.00	1.86	2.78	7.00	2.82	2.14	4.00
	IPLS_Stepwise_50	1.38	1.98	5.00	0.99	1.27	11.00	0.63	0.45	3.00	**2.11**	**1.86**	**4.00**	2.46	1.84	7.00
	IPLS_Stepwise_100	0.63	1.69	13.00	1.64	1.59	6.00	**0.32**	**0.52**	**12.00**	1.87	2.54	6.00	2.48	1.76	4.00
First derivative	*No variable selection*	0.88	1.62	7.00	1.24	1.66	6.00	0.21	0.58	15.00	1.53	2.38	9.00	0.94	2.62	18.00
	IPLS_Backward_10	1.13	1.53	6.00	1.37	1.59	9.00	0.61	0.53	3.00	2.67	3.09	4.00	2.69	2.66	4.00
	IPLS_Backward_25	2.38	2.16	2.00	2.12	2.40	3.00	0.66	0.53	1.00	1.63	2.06	12.00	3.55	3.15	3.00
	IPLS_Backward_50	1.44	1.96	10.00	2.13	2.21	4.00	0.63	0.48	2.00	2.52	2.49	6.00	2.96	3.53	6.00
	IPLS_Backward_100	1.47	1.31	3.00	1.84	1.27	3.00	0.62	0.45	5.00	2.29	2.39	7.00	2.21	2.19	9.00
	IPLS_Stepwise_10	0.71	1.87	10.00	1.25	1.78	9.00	0.54	0.48	6.00	2.34	2.57	8.00	2.19	2.12	10.00
	IPLS_Stepwise_25	1.14	1.93	10.00	0.99	1.48	11.00	0.62	0.45	3.00	2.24	3.68	5.00	2.66	2.58	7.00
	IPLS_Stepwise_50	1.05	1.68	9.00	1.13	1.15	9.00	0.51	0.54	5.00	2.16	2.72	4.00	**1.98**	**1.78**	**5.00**
	IPLS_Stepwise_100	1.14	1.27	5.00	**1.04**	**1.10**	**9.00**	0.35	0.53	10.00	1.92	2.28	7.00	2.03	1.64	6.00
Second derivative	*No variable selection*	0.86	1.62	8.00	1.01	1.20	9.00	0.21	0.55	14.00	0.29	3.07	20.00	0.66	2.65	19.00
	IPLS_Backward_10	1.49	2.12	4.00	1.62	1.80	6.00	0.61	0.46	4.00	2.26	3.29	5.00	2.50	1.62	3.00
	IPLS_Backward_25	1.38	1.81	7.00	2.54	2.27	2.00	0.62	0.49	2.00	2.20	2.84	5.00	2.47	1.79	3.00
	IPLS_Backward_50	1.39	1.65	5.00	1.35	1.26	6.00	0.66	0.47	2.00	2.76	2.86	1.00	2.72	2.45	5.00
	IPLS_Backward_100	1.55	1.87	5.00	1.86	1.45	3.00	0.62	0.47	3.00	1.95	2.11	7.00	2.92	2.19	3.00
	IPLS_Stepwise_10	1.21	1.66	7.00	1.25	1.82	11.00	0.57	0.60	6.00	1.93	2.94	9.00	1.81	2.65	12.00
	IPLS_Stepwise_25	1.39	1.60	3.00	1.00	1.75	13.00	0.63	0.47	1.00	1.56	3.45	10.00	2.40	2.17	6.00
	IPLS_Stepwise_50	1.13	1.67	5.00	1.12	1.14	7.00	0.61	0.50	2.00	1.84	3.05	9.00	2.36	1.41	4.00
	IPLS_Stepwise_100	1.31	1.36	5.00	1.17	1.11	5.00	0.65	0.53	1.00	1.85	2.44	4.00	2.12	1.86	7.00

References

Andersen, C.M., Bro, R., 2010. Variable selection in regression—a tutorial. Journal of Chemometrics 24, 728-737.

Chen, X., Wu, J., Zhou, S., Yang, Y., Ni, X., Yang, J., Zhu, Z., Shi, C., 2009. Application of near-infrared reflectance spectroscopy to evaluate the lutein and β-carotene in Chinese kale. Journal of food composition and analysis 22, 148-153.

Cornillon, P.A., 2010. Statistiques avec R: 2e édition augmentée. Presses universitaires de Rennes.

Cortés, V., Ortiz, C., Aleixos, N., Blasco, J., Cubero, S., Talens, P., 2016. A new internal quality index for mango and its prediction by external visible and near-infrared reflection spectroscopy. Postharvest Biology and Technology 118, 148-158.

Etienne, A., Génard, M., Bancel, D., Benoit, S., Bugaud, C., 2013. A model approach revealed the relationship between banana pulp acidity and composition during growth and post harvest ripening. Scientia Horticulturae 162, 125-134.

Génard, M., Gouble, B., 2005. ETHY. A Theory of Fruit Climacteric Ethylene Emission. Plant Physiology 139, 531-545.

Jha, S.N., Jaiswal, P., Narsaiah, K., Gupta, M., Bhardwaj, R., Singh, A.K., 2012a. Non-destructive prediction of sweetness of intact mango using near infrared spectroscopy. Scientia Horticulturae 138, 171-175.

Jha, S.N., Jaiswal, P., Narsaiah, K., Gupta, M., Bhardwaj, R., Singh, A.K., 2012b. Non-destructive prediction of sweetness of intact mango using near infrared spectroscopy. Scientia Horticulturae 138, 171-175.

Jha, S.N., Kingsly, A.R.P., Chopra, S., 2006. Non-destructive determination of firmness and yellowness of mango during growth and storage using visual spectroscopy. Biosystems Engineering 94, 397-402.

Jha, S.N., Narsaiah, K., Jaiswal, P., Bhardwaj, R., Gupta, M., Kumar, R., Sharma, R., 2014. Nondestructive prediction of maturity of mango using near infrared spectroscopy. Journal of Food Engineering 124, 152-157.

Joas, J., Caro, Y., Lechaudel, M., 2009. Comparison of postharvest changes in mango (cv Cogshall) using a Ripening class index (Rci) for different carbon supplies and harvest dates. Postharvest biology and technology 54, 25-31.

Joas, J., Vulcain, E., Desvignes, C., Morales, E., Léchaudel, M., 2012. Physiological age at harvest regulates the variability in postharvest ripening, sensory and nutritional characteristics of mango (Mangifera indica L.) cv. Coghshall due to growing conditions. Journal of the Science of Food and Agriculture 92, 1282-1290.

Joas, J., Vulcain, E., Léchaudel, M., 2010. Effect of fruit position in the canopy on physiological age and physicochemical composition of mango'Cogshall', IX International Mango Symposium 992, pp. 123-128.

Jordan, R.B., Walton, E.F., Klages, K.U., Seelye, R.J., 2000. Postharvest fruit density as an indicator of dry matter and ripened soluble solids of kiwifruit. Postharvest Biology and Technology 20, 163-173.

Kapse, B.M., Katrodia, J., 1996. Ripening behaviour of'Kesar'mangoes in relation to specific gravity, V International Mango Symposium 455, pp. 669-678.

Lê, S., Josse, J., Husson, F., 2008. FactoMineR: an R package for multivariate analysis. Journal of statistical software 25, 1-18.

Lebrun, M., Plotto, A., Goodner, K., Ducamp, M.-N., Baldwin, E., 2008. Discrimination of mango fruit maturity by volatiles using the electronic nose and gas chromatography. Postharvest Biology and Technology 48, 122-131.

Léchaudel, M., Joas, J., 2007. An overview of preharvest factors influencing mango fruit growth, quality and postharvest behaviour. Brazilian Journal of Plant Physiology 19, 287-298.

Léchaudel, M., Joas, J., Caro, Y., Génard, M., Jannoyer, M., 2005. Leaf:fruit ratio and irrigation supply affect seasonal changes in minerals, organic acids and sugars of mango fruit. Journal of the Science of Food and Agriculture 85, 251-260.

Lechaudel, M., Urban, L., Joas, J., 2010. Chlorophyll Fluorescence, a Nondestructive Method To Assess Maturity of Mango Fruits (Cv. 'Cogshall') without Growth Conditions Bias. Journal of Agricultural and Food Chemistry 58, 7532-7538.
Léchaudel, M., Vercambre, G., Lescourret, F., Normand, F., Génard, M., 2007. An analysis of elastic and plastic fruit growth of mango in response to various assimilate supplies. Tree Physiology 27, 219-230.
Lescourret, F., Moitrier, N., Valsesia, P., Génard, M., 2011. QualiTree, a virtual fruit tree to study the management of fruit quality. I. Model development. Trees 25, 519-530.
Litz, R.E., 2009. The Mango, 2nd Edition: Botany, Production and Uses. Wiley Online Library.
Liu, H.-F., Génard, M., Guichard, S., Bertin, N., 2007. Model-assisted analysis of tomato fruit growth in relation to carbon and water fluxes. Journal of Experimental Botany 58, 3567-3580.
Lobit, P., Génard, M., Wu, B.H., Soing, P., Habib, R., 2003. Modelling citrate metabolism in fruits: responses to growth and temperature. Journal of Experimental Botany 54, 2489-2501.
Lumley, T., Miller, A., 2009. Leaps: regression subset selection. R package version 2.9. Online at http://CRAN. R-project. org/package= leaps.
Maniwara, P., Nakano, K., Boonyakiat, D., Ohashi, S., Hiroi, M., Tohyama, T., 2014. The use of visible and near infrared spectroscopy for evaluating passion fruit postharvest quality. Journal of Food Engineering 143, 33-43.
Marques, E.J.N., de Freitas, S.T., Pimentel, M.F., Pasquini, C., 2016. Rapid and non-destructive determination of quality parameters in the 'Tommy Atkins' mango using a novel handheld near infrared spectrometer. Food Chemistry 197, 1207-1214.
McAtee, P., Karim, S., Schaffer, R., David, K., 2013. A dynamic interplay between phytohormones is required for fruit development, maturation, and ripening. Frontiers in plant science 4, 79.
McGlone, V.A., Jordan, R.B., Martinsen, P.J., 2002a. Vis/NIR estimation at harvest of pre-and post-storage quality indices for 'Royal Gala'apple. Postharvest Biology and Technology 25, 135-144.
McGlone, V.A., Jordan, R.B., Seelye, R., Martinsen, P.J., 2002b. Comparing density and NIR methods for measurement of Kiwifruit dry matter and soluble solids content. Postharvest Biology and Technology 26, 191-198.
Medlicott, A.P., Bhogal, M., Reynolds, S.B., 1986. Changes in peel pigmentation during ripening of mango fruit (Mangifera indica var. Tommy Atkins). Annals of Applied Biology 109, 651-656.
Medlicott, A.P., Thompson, A.K., 1985. Analysis of sugars and organic acids in ripening mango fruits (Mangifera indica L. var Keitt) by high performance liquid chromatography. Journal of the Science of Food and Agriculture 36, 561-566.
Mevik, B.-H., Wehrens, R., 2007. The pls package: principal component and partial least squares regression in R. Journal of Statistical Software 18, 1-24.
Nagle, M., Mahayothee, B., Rungpichayapichet, P., Janjai, S., Müller, J., 2010. Effect of irrigation on near-infrared (NIR) based prediction of mango maturity. Scientia Horticulturae 125, 771-774.
Nicolaï, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K.I., Lammertyn, J., 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. Postharvest Biology and Technology 46, 99-118.
Nordey, T., Joas, J., Davrieux, F., Chillet, M., Léchaudel, M., 2017. Robust NIRS models for non-destructive prediction of mango internal quality. Scientia Horticulturae 216, 51-57.
Nordey, T., Léchaudel, M., Génard, M., Joas, J., 2014. Spatial and temporal variations in mango colour, acidity, and sweetness in relation to temperature and ethylene gradients within the fruit. Journal of plant physiology 171, 1555-1563.
Nordey, T., Léchaudel, M., Génard, M., Joas, J., 2016. Factors affecting ethylene and carbon dioxide concentrations during ripening: Incidence on final dry matter, total soluble solids content and acidity of mango fruit. Journal of Plant Physiology 196, 70-78.
Palmer, J.W., Harker, F.R., Tustin, D.S., Johnston, J., 2010. Fruit dry matter concentration: a new quality metric for apples. Journal of the Science of Food and Agriculture 90, 2586-2594.

Poiroux-Gonord, F., Fanciullino, A.L., Bert, L., Urban, L., 2012. Effect of fruit load on maturity and carotenoid content of clementine (Citrus clementina Hort. ex Tan.) fruits. Journal of the Science of Food and Agriculture 92, 2076-2083.
R Development Core Team, 2012. R: A Language and Environment for Statistical Computing.
Rosalie, R., Joas, J., Deytieux-Belleau, C., Vulcain, E., Payet, B., Dufossé, L., Léchaudel, M., 2015. Antioxidant and enzymatic responses to oxidative stress induced by pre-harvest water supply reduction and ripening on mango (Mangifera indica L. cv. 'Cogshall') in relation to carotenoid content. Journal of Plant Physiology 184, 68-78.
Rungpichayapichet, P., Mahayothee, B., Khuwijitjaru, P., Nagle, M., Muller, J., 2015. Non-destructive determination of beta-carotene content in mango by near-infrared spectroscopy compared with colorimetric measurements. Journal of Food Composition and Analysis 38, 32-41.
Rungpichayapichet, P., Mahayothee, B., Nagle, M., Khuwijitjaru, P., Müller, J., 2016. Robust NIRS models for non-destructive prediction of postharvest fruit ripeness and quality in mango. Postharvest Biology and Technology 111, 31-40.
Saranwong, S., Sornsrivichai, J., Kawano, S., 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, 137-145.
Schmilovitch, Z.e., Mizrach, A., Hoffman, A., Egozi, H., Fuchs, Y., 2000. Determination of mango physiological indices by near-infrared spectrometry. Postharvest Biology and Technology 19, 245-252.
Song, J., Deng, W., Beaudry, R.M., Armstrong, P.R., 1997. Changes in Chlorophyll Fluorescence of Apple Fruit during Maturation, Ripening, and Senescence. HortScience 32, 891-896.
Stevens, A., Ramirez-Lopez, L., 2013. Miscellaneous functions for processing and sample selection of vis-NIR diffuse reflectance data.
Subedi, P., Walsh, K.B., Owens, G., 2007. Prediction of mango eating quality at harvest using short-wave near infrared spectrometry. Postharvest Biology and Technology 43, 326-334.
Urbano Bron, I., Vasconcelos Ribeiro, R., Azzolini, M., Pedro Jacomino, A., Caruso Machado, E., 2004. Chlorophyll fluorescence as a tool to evaluate the ripening of 'Golden' papaya fruit. Postharvest Biology and Technology 33, 163-173.
Vasquez-Caicedo, A.L., Sruamsiri, P., Carle, R., Neidhart, S., 2005. Accumulation of All-trans-β-carotene and Its 9-cis and 13-cis Stereoisomers during Postharvest Ripening of Nine Thai Mango Cultivars. Journal of Agricultural and Food Chemistry 53, 4827-4835.