




## Investigation of the internal quality of ‘Palmer’ and ‘Tommy Atkins’ mangoes by near-infrared spectroscopy

Thamires Gomes ANTUNES<sup>1</sup> , Daniela Santana de SOUZA<sup>1</sup> , Bruna Ariel Dias GUARIGLIA<sup>1</sup> , Diogo Pedrosa Corrêa da SILVA<sup>1</sup> , Cristiane Maria Ascari MORGADO<sup>2\*</sup> , Lucidarcy Martins da MATTA<sup>1</sup> , Gilmarcos de Carvalho CORRÊA<sup>1</sup> , Luis Carlos CUNHA JÚNIOR<sup>1</sup> 

### Abstract

This study focuses on optimizing mango harvesting and minimizing waste by using a handheld near-infrared (NIR) spectrophotometer to develop predictive models for assessing the quality of ‘Palmer’ and ‘Tommy Atkins’ mangoes. It aims to enhance mechanical resistance and reduce post-harvest losses. The study created prediction models for key quality attributes: soluble solids (SS), titratable acidity (TA), and dry matter (DM). For ‘Palmer’ mangoes, the first derivative of Savitzky–Golay (SG1) yielded the best SS predictions (the coefficient of determination for prediction [ $R^2_p$ ] = 0.69, the square root of the mean error of prediction [ $RMSE_p$ ] = 1.56%, and the standard deviation ratio of prediction [ $SDR_p$ ] = 1.80). For ‘Tommy Atkins’ mangoes, the second derivative of Savitzky–Golay (SG2) was more effective ( $R^2_p$  = 0.72,  $RMSE_p$  = 2.43%, and  $SDR_p$  = 1.85). TA prediction models showed that SG2 was more effective for ‘Palmer’ ( $R^2_p$  = 0.56,  $RMSE_p$  = 0.14%, and  $SDR_p$  = 1.48), while multiplicative signal correction pre-treatment worked better for ‘Tommy Atkins’ ( $R^2_p$  = 0.59,  $RMSE_p$  = 0.13%, and  $SDR_p$  = 1.55). For DM predictions, SG1 was optimal for ‘Palmer’ ( $R^2_p$  = 0.83,  $RMSE_p$  = 0.95, and  $SDR_p$  = 2.44) and SG2 for ‘Tommy Atkins’ ( $R^2_p$  = 0.79,  $RMSE_p$  = 1.36, and  $SDR_p$  = 2.00). In conclusion, the handheld NIR spectrophotometer shows promise for accurate quality assessments in the mango production chain, enabling better decision-making on harvest timing and reducing post-harvest losses.

**Keywords:** *Mangifera indica* L; soluble solids; titratable acidity; dry matter; short-wave near-infrared region; near-infrared spectroscopy.

**Practical Application:** Near-infrared technology optimizes mango harvest and reduces losses through precise assessments.

## 1 INTRODUCTION

Mango (*Mangifera indica* L) is one of the most widely cultivated tropical fruits globally, known for its vibrant color, refreshing taste, and nutritional and therapeutic values. In 2021, India was the largest producer of mangoes, followed by China, Indonesia, Pakistan, Mexico, and Brazil (Lauricella et al., 2017; Maldonado-Celis et al., 2019; Zhang et al., 2022). That same year, mango became Brazil’s top fruit export, with over 272 million tons shipped, generating more than US\$ 1.218 billion in revenue (MAPA, 2023).

Consumers worldwide often choose fruits based on criteria such as skin integrity, color, and pulp firmness, which also apply to mango selection for its appealing taste. A study on the relationship between the dry matter (DM) content and high consumer acceptance scores of ripe ‘Palmer’ and ‘Tommy Atkins’ mangoes from Brazilian orchards found acceptance scores of 7–9 (“like moderately”) for minimum DM values of

137 and 144 g kg<sup>-1</sup>, respectively (Freitas et al., 2022). In Italy, research showed that 80% of evaluators rated mangoes harvested after 119 days from full flowering as “good” or “excellent,” correlating with SS and DM contents exceeding 12% each (Gianguzzi et al., 2021).

Harvest time plays a crucial role in determining the internal quality of mangoes intended for industry or consumer markets. As a climacteric fruit, mangoes can be harvested when physiologically mature and stored under optimal conditions to slow ripening, depending on their final destination (Yahia, 2019). Harvest decisions are based on field assessments (such as peel color and texture) combined with destructive physicochemical analyses (Lobo & Sidhu, 2017). The key indicators of mango maturity include soluble solids (SS), DM, and titratable acidity (TA) contents (Subedi et al., 2007). However, the most reliable parameter for determining harvest time remains the full flowering period (126–133 days) (Gianguzzi et al., 2021).

Received: Oct. 26, 2024.

Accepted: Nov. 19, 2024.

<sup>1</sup>Universidade Federal de Goiás, Samambaia Campus, School of Agronomy, Goiânia, GO, Brazil.

<sup>2</sup>Universidade Estadual de Goiás, Southwest Campus, Quirinópolis, GO, Brazil.

\*Corresponding author: cristiane.morgado@ueg.br

Conflict of interest: nothing to declare.

Funding: Goiás State Research Support Foundation (FAPEG).

Near-infrared (NIR) spectroscopy is a more precise and consistent method for estimating fruit characteristics than relying on flowering time. Flowering can be irregular, with multiple flowering events in the same orchard, causing variations that complicate uniform fruit analysis. Environmental factors and cultivation practices further affect flowering regularity and intensity. In contrast, NIR spectroscopy directly measures properties such as SS and TA, offering precise and repeatable assessments without the variability of flowering time. This analytical technique allows for immediate and direct fruit quality monitoring, making it a more reliable approach for agricultural evaluations (Mishra & Woltering, 2023; Nicolai et al., 2007).

Advancements in computational technology have enabled non-destructive methods to accurately determine the optimal harvest time for fruits. NIR spectroscopy stands out as a rapid, non-destructive, low-cost analytical technique that requires minimal sample preparation and no reagents. It has shown promise in assessing harvest indicators for various fruits, including bananas (Subedi & Walsh, 2011), cherries (Li et al., 2018), apples (Fan et al., 2020), pears (Yu & Yao, 2022), tomatoes (Brito et al., 2022), and mangoes (Anderson et al., 2020; Freitas et al., 2022; Munawar et al., 2022; Shah et al., 2021). This method is particularly valuable due to the absence of a universal maturity pattern among mango cultivars, which varies with different growing conditions (Yahia, 2019).

Despite the development of accurate calibration and prediction models using NIR spectroscopy for 'Palmer' (Santos Neto et al., 2017) and 'Tommy Atkins' (Marques et al., 2016) mangoes in Brazil, harvest indicators can vary regionally. To address future demands for regional calibration models, the internal quality of 'Tommy Atkins' and 'Palmer' mangoes was studied using chemometric models. These models, applied through NIR spectroscopy, were developed to predict key quality indicators such as SS, TA, and DM, ensuring accurate assessments tailored to different Brazilian regions.

## 2 MATERIAL AND METHODS

### 2.1 Materials

'Palmer' and 'Tommy Atkins' mangoes (*M. indica*) were harvested from farms located in Campestre (16°44'17" latitude South, 49°41'48" longitude West, 693 m altitude), Silvânia (16°38'35" latitude South, 48°36'15" longitude West, 877 m altitude), and Cristalina (16°46'4" latitude South, 47°36'47" longitude West, 1237 m altitude), Goiás, Brazil. Employing the size and shape of commercial fruits as a reference, the fruit was harvested at randomized developmental stages and subsequently transported in plastic boxes at an ambient temperature. The selection criteria ( $n_{\text{Palmer}} = 256$ , and  $n_{\text{Tommy Atkins}} = 207$ ) were based on surface quality attributes (absence of physical damage and diseases), visually classified for peel color similarity, and stored at 25°C until another evaluation.

### 2.2 Analytical methodology

The mangoes underwent destructive (reference) and non-destructive techniques, evaluating SS, TA, and DM. Two

locations were taken from each fruit (Figure 1) from both sides of the equatorial region, equidistant from proximal and distal ends (Subedi et al., 2007).

### 2.3 Spectral achievement

NIR spectra were collected using the F-750 portable NIR spectrometer (Felix Instruments, Washington, USA) within a wavelength range of 250–1100 nm, employing interreflectance geometry with a resolution of 3 nm. For each marked location, one spectrum was obtained, in the equatorial region, equidistant proximally and distally (Figure 1). Each spectrum was the result of an average of four scans to enhance the signal-to-noise ratio. The resulting spectra were then split between training and test sets for subsequent analysis (Subedi et al., 2007).

### 2.4 Reference evaluations

#### 2.4.1 Soluble solids content

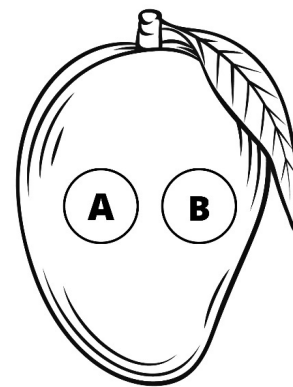
The SS content was determined using a digital refractometer (Reichert Brix/RI-Check, NY, USA) (met. 932.12) (AOAC, 1997). After peel removal, the mango pulp was macerated, and juice was obtained through gauze, with results expressed in percentage (%).

#### 2.4.2 Titratable acidity content

The TA content was obtained through volumetric titration (met. 942.15) (AOAC, 1997). Subsequently, 5 g of the pulp juice and 50 mL of distilled water were mixed under magnetic stirring. Sodium hydroxide (100 mM) was titrated under constant agitation using 1% phenolphthalein alcoholic solution as an indicator, with results expressed in percentage of citric acid.

#### 2.4.3 Dry matter content

The DM was determined via gravimetric methodology. After removing the 1–2-mm-thick mango peel, a cylindrical sample portion of 27 mm in diameter and 10 mm in depth was subjected to a forced air circulation oven at 105°C until



**Figure 1.** Schematic diagram of the marked places for near-infrared spectral and reference analysis in the sampling of mangoes.

reaching a constant weight (Subedi et al., 2007). The results were expressed as percentage (%).

## 2.5 Chemometrics

The Unscrambler® software version 10.0.3 (CAMO, Oslo, Norway) was employed for the analysis of spectral data and reference values. The data, with each sample representing an individual fruit, underwent principal component analysis (PCA) to reduce dimensionality and identify key variables for fruit discrimination. Following this, the Kennard–Stone algorithm was applied to divide the fruits into calibration (two-thirds of the samples) and prediction (one-third of the samples) sets. The division was carefully conducted to ensure that fruits included in the calibration set were distinct from those in the prediction set, thereby preventing any overlap between the two sets. This rigorous approach guarantees that the calibration and prediction sets are independent, enhancing the robustness and validity of the statistical analysis. The spectra were pre-processed using multiplicative signal correction (MSC) and the second polynomial order of the first and second derivatives of Savitzky–Golay (SG1 and SG2, respectively) with smoothing windows of either 5 points (2+2) or 7 points (5+5). Calibration models were developed utilizing partial least squares (PLS) regression with the non-linear iterative partial least squares algorithm. Cross-validation involved five segments with random sample selection.

Models' performance was evaluated through the square root of the mean error of calibration ( $RMSE_C$ ), the square root of the mean error of cross-validation ( $RMSE_{cv}$ ), the square root of the mean error of prediction ( $RMSE_p$ ), the standard error of prediction (SEP), the coefficient of determination for calibration ( $R^2_c$ ), the coefficient of determination for cross-validation ( $R^2_{cv}$ ), the coefficient of determination for prediction ( $R^2_p$ ), and standard deviation ratio (SDR) (Golic & Walsh, 2006; Toscano et al., 2017).

## 3 RESULTS AND DISCUSSION

### 3.1 Reference evaluations: soluble solids, titratable acidity, and dry matter

Table 1 shows the calibration sets for 'Palmer' and 'Tommy Atkins' mangoes, displaying average SS values of 11.17 and 13.59%, respectively. The SS content of the Palmer cultivar ranged from 5.3 to 19.6%, a range consistent with findings from researchers using 'Palmer' mangoes from São Paulo, Brazil, via NIR spectroscopy (Santos Neto et al., 2017). The Tommy Atkins cultivar (Table 1) showed a range of SS between 4.3 and 22.78%, with variation values akin to those documented from 'Tommy Atkins' mangoes harvested from Bahia, Brazil. Physiologically mature fruits have a 7–9% range, a maximum of 20% at the point of maturation completion, attributed to the biosynthetic process and degradation of polysaccharides accumulated during fruit growth. Consequently, the sampled population of 'Palmer' and 'Tommy Atkins' mangoes predominantly comprised mature fruits, encompassing stages from immaturity to full maturation, reflecting the entire developmental spectrum.

The calibration set for TA in 'Palmer' and 'Tommy Atkins' mangoes revealed average values of 0.51% citric acid and 0.50% citric acid, respectively (Table 1). 'Palmer' mangoes exhibited a TA range of 0.19–0.98% citric acid, while 'Tommy Atkins' mangoes displayed a TA range of 0.04–1.12% citric acid, i.e., most of the fruit populations are ripe. The researchers observed no significant difference (Tukey's test, 5%) between the TA values of the 'Palmer' and 'Tommy Atkins' mangoes harvested in summer and winter: the TA ranges were consistent, measuring 1.1–1.2% for 'Palmer' and 1.0–1.2% for 'Tommy Atkins' mangoes (Freitas et al., 2022). The TA content in mangoes can be mainly attributed to the levels of citric and malic acids, which diminish during fruit development due to their utilization as a substrate in respiration or their conversion into sugars, enhancing pulp stability, flavor, and aroma (Yahia, 2019).

**Table 1.** Descriptive statistics of calibration and prediction for soluble solids, titratable acidity, and dry matter in 'Palmer' and 'Tommy Atkins' mangoes.

Soluble solids (%)								
Variety	N	Calibration			N	Prediction		
		Mean	Range	SD		Mean	Range	SD
Palmer	104	11.17	5.9–19.6	2.97	51	10.00	5.30–16.50	2.80
Tommy Atkins	67	13.59	4.3–22.78	4.59	33	13.73	6.30–21.51	4.34
Titratable acidity (% citric acid)								
Variety	N	Calibration			N	Prediction		
		Mean	Range	SD		Mean	Range	SD
Palmer	77	0.51	0.19–0.98	0.18	38	0.51	0.15–0.98	0.19
Tommy Atkins	80	0.50	0.04–1.12	0.21	40	0.51	0.15–0.92	0.20
Dry matter (%)								
Variety	N	Calibration			N	Prediction		
		Mean	Range	SD		Mean	Range	SD
Palmer	166	19.77	15.72–38.11	2.71	82	20.08	16.74–36.43	2.64
Tommy Atkins	116	15.58	8.78–23.70	3.77	57	15.54	9.56–20.84	3.00

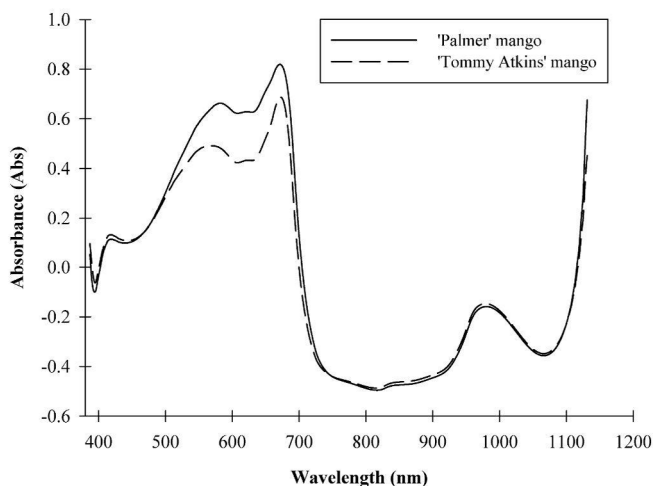
SD: standard deviation.

The prediction sets for 'Palmer' and 'Tommy Atkins' mangoes revealed average SS values of  $10.00\% \pm 2.80\%$  and  $13.73\% \pm 4.34\%$ , respectively (Table 1). The SS range for 'Palmer' was 5.3–16.50%, while 'Tommy Atkins' exhibited a range of 6.30–21.51%. The average TA values in the prediction set for 'Palmer' and 'Tommy Atkins' mangoes were  $0.51\% \pm 0.19\%$  citric acid and  $0.51\% \pm 0.20\%$  citric acid, respectively (Table 1). The TA value of 'Palmer' ranged from 0.15 to 0.98% citric acid and that of 'Tommy Atkins' ranged from 0.15 to 0.92% citric acid. The average DM values were  $20.08\% \pm 2.64\%$  for 'Palmer' and  $15.54\% \pm 3.00\%$  for 'Tommy Atkins' mangoes, obtained from the prediction set (Table 1). The DM value of 'Palmer' ranged from 16.74 to 36.43%, while that of 'Tommy Atkins' ranged from 9.56 to 20.84%.

### 3.2 Chemometric assessments

Figure 2 shows the mean raw spectra of 'Palmer' and 'Tommy Atkins' mangoes acquired using a portable spectrometer. The initial absorption band of the spectrum aligns with the end of the visible wavelength region (387–750 nm) with heightened absorptions in violet (380–440 nm), green (500–565 nm), and red (625–740 nm) spectral regions. This characterizes typical color variations in the skins of 'Palmer' and 'Tommy Atkins' mangoes. Notably, both spectra feature a prominent peak at 672 nm, within the red color range, signifying the presence of red pigments including anthocyanins, active compounds in ripe fruits. The short-wave near-infrared region (SWNIR) (Subedi et al., 2007) from 750 to 1,131 nm reveals a high absorption band in the 957–1,014 nm range corresponding to the molecular vibrations of water. This aligns with the high moisture content in 'Palmer' and 'Tommy Atkins' mangoes, measured at 79.7 and 85.8%, respectively.

Before regression analysis, specific spectral ranges were chosen for each quality attribute to evaluate spectral data information through PCA (Marques et al., 2016). The ranges 699–999 nm (Santos Neto et al., 2017), 743–903 nm (Subedi et al., 2012), and 699–981 nm (Santos Neto et al., 2017) were selected for



**Figure 2.** Mean raw near-infrared spectra of 'Palmer' and 'Tommy Atkins' mangoes.

SS, TA, and DM, respectively, in 'Palmer' and 'Tommy Atkins' mangoes. PCA models for SS (Figure 3), TA (Figure 4), and DM (Figure 5) from the best treatment options explained 88–100% of data variability, with PC1 and PC2 accounting for 97–100% of the variance.

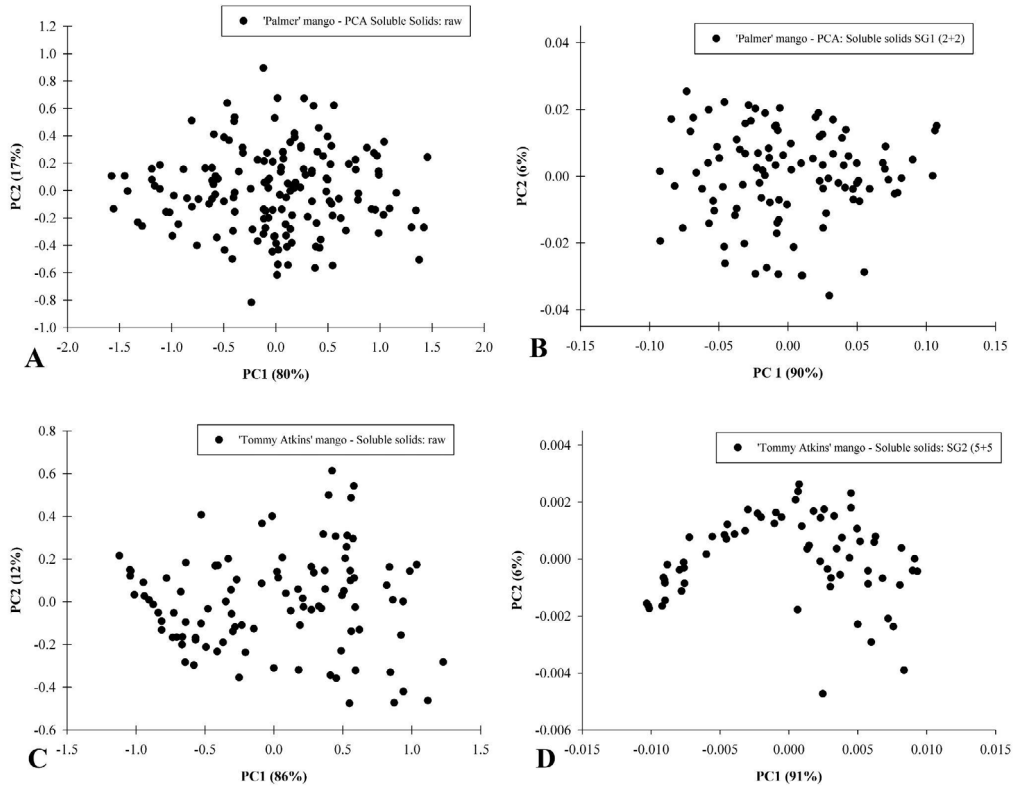
After spectral transformation (pre-treatment) and exploratory and regression analysis of each quality attribute, the best prediction results for SS, TA, and DM using SG1 (2+2), SG2 (5+5), and SG1 (5+5), respectively, were achieved for 'Palmer' mangoes (Figure 6). In the case of 'Tommy Atkins' mangoes, the best prediction results for the same sequence of attributes were obtained using SG2 (5+5), MSC, and SG2 (2+2), respectively (Figure 7).

### 3.3 Soluble solids content

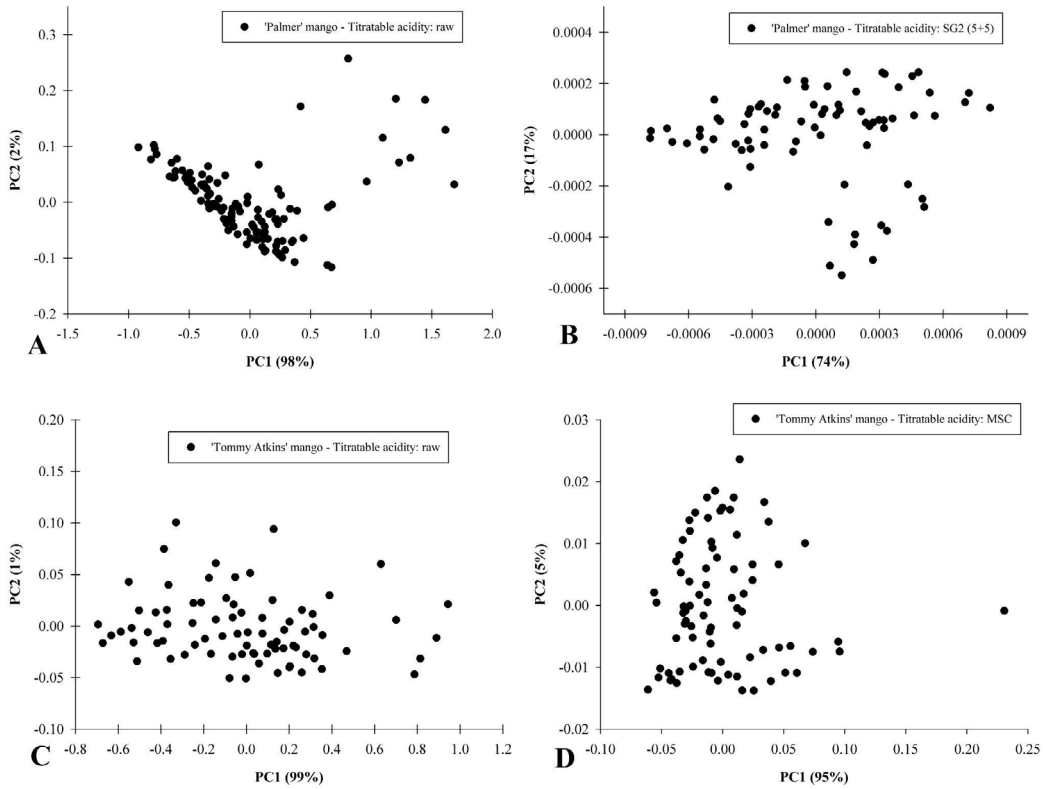
For 'Palmer' mangoes, SG1 with a 2+2 smoothing window in the 699–999 nm spectral window yielded superior SS prediction results ( $R^2_p = 0.69$ ,  $RMSE_p = 1.56\%$ , and  $SDR_p = 1.80$ ), generating a calibration model employing 12 latent variables (LV),  $R^2_{cv} = 0.74$ , and  $RMSE_{cv} = 1.97\%$  (Table 2). The 12 LVs were selected based on a systematic cross-validation process using the modeling algorithm available in the Unscrambler software. The criteria for selecting these variables were based on balancing the model fit (measured by the coefficient of determination,  $R^2$ ) and the root mean square error of calibration ( $RMSE_c$ ). Although 12 LVs might seem high for a dataset with hundreds of fruits, this number was deemed optimal by the algorithm because reducing the number of LVs led to a significant decrease in  $R^2$  and an increase in  $RMSE_c$ , indicating a loss in the precision and reliability of the predictive model. Therefore, the choice of 12 LVs reflects the need to maintain sufficient complexity to capture the relevant variations in the data, ensuring robust and high-quality prediction.

For 'Tommy Atkins' mangoes, SG2 with a 5+5 smoothing window in the 699–999 nm spectral window yielded enhanced SS prediction results ( $R^2_p = 0.72$ ,  $RMSE_p = 2.43\%$ , and  $SDR_p = 1.85$ ), generating a calibration model with seven LVs,  $R^2_{cv} = 0.89$ , and  $RMSE_{cv} = 1.73\%$  (Table 2). Also called RPD by other researchers, as proposed by Nicolai et al. (2007), the SDR parameter was employed to assess models' performance regarding sample variation. Specifically, the values of  $1.5 < SDR < 2$  suggest that the model can distinguish low from high response variables, the values of  $2 < SDR < 2.5$  denote ordinary precision, and the values of  $2.5 < SDR < \text{upper}$  correspond to 'good' and/or 'excellent' accuracy.

Nicolai et al. (2007) explained that the RPD is defined as the ratio of the standard deviation of the response variable to the  $RMSE_p$  or  $RMSE_{cv}$  (some authors use the term SDR). Standard deviation ratio and ratio of performance to deviation are terms used interchangeably to describe the same predictive model's evaluation metric. SDR is defined as the ratio of the standard deviation of the response variable to the  $RMSE_p$  or the  $RMSE_{cv}$ . Similarly, RPD is described as the ratio of the standard deviation of the response variable to the  $RMSE_p$  or  $RMSE_{cv}$ . Although the term RPD is more commonly used in some publications, both terms refer to the same concept of assessing model accuracy



**Figure 3.** (A) The principal component analysis results of ‘Palmer’ mangoes’ soluble solids in relative absorbance and (B) with the application of the Savitzky–Golay first derivative (2+2); (C) the principal component analysis results of ‘Tommy Atkins’ mangoes’ soluble solids in relative absorbance and (D) with the application of the Savitzky–Golay second derivative (5+5).



**Figure 4.** (A) The principal component analysis results of ‘Palmer’ mangoes’ titratable acidity in relative absorbance and (B) with the application of the Savitzky–Golay second derivative (5+5); (C) the principal component analysis results of ‘Tommy Atkins’ mangoes’ titratable acidity in relative absorbance and (D) with the application of the multiplicative signal correction.

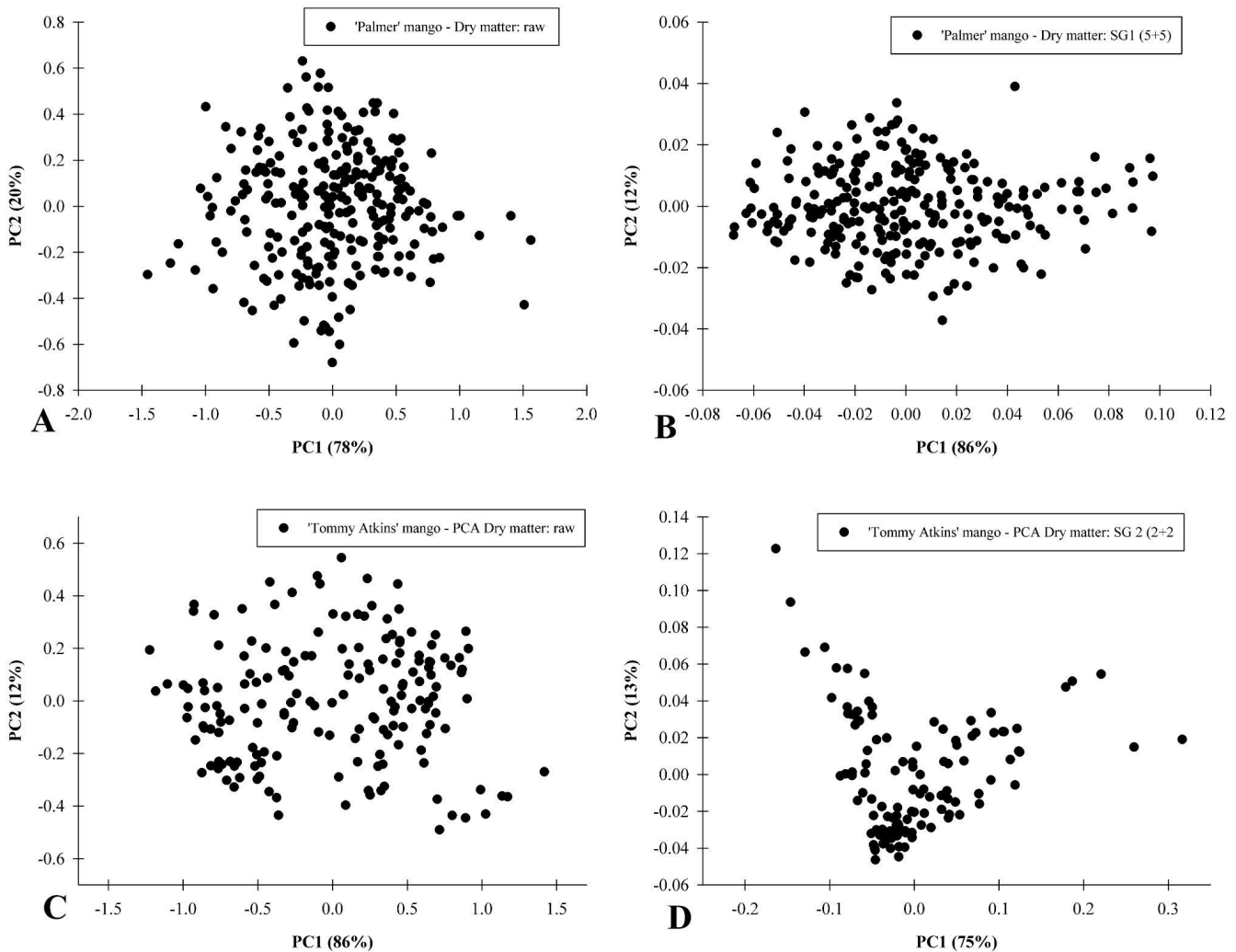
and robustness. Therefore, SDR and RPD can be considered synonymous in the literature and within this study.

The  $RMSE_{cv}$  value (1.97%) for SS in 'Palmer' mangoes surpasses that reported by Santos Neto et al. (2017) while utilizing a portable instrument (F-750 Felix Instruments), Standard Normal Variate (SNV), and SG1 in the spectral range of 699–999 nm, where they reported values of  $R^2_{cv} = 0.87$ ,  $RMSE_{cv} = 1.39\%$ , and  $SDR = 2.78$  for 'Palmer' mangoes. Marques et al. (2016) while operating MicroNIR 1700 spectrometer and using SNV pre-processing in the spectral window of 1,200–2,400 nm obtained  $R^2_{cv} = 0.83$  and  $RMSE_{cv} = 0.88\%$  for the SS of 'Tommy Atkins' mangoes, values lower than those obtained in the present study ( $R^2_{cv} = 0.89$  and  $RMSE_{cv} = 1.73\%$ ) for the SS of 'Tommy Atkins' mangoes.

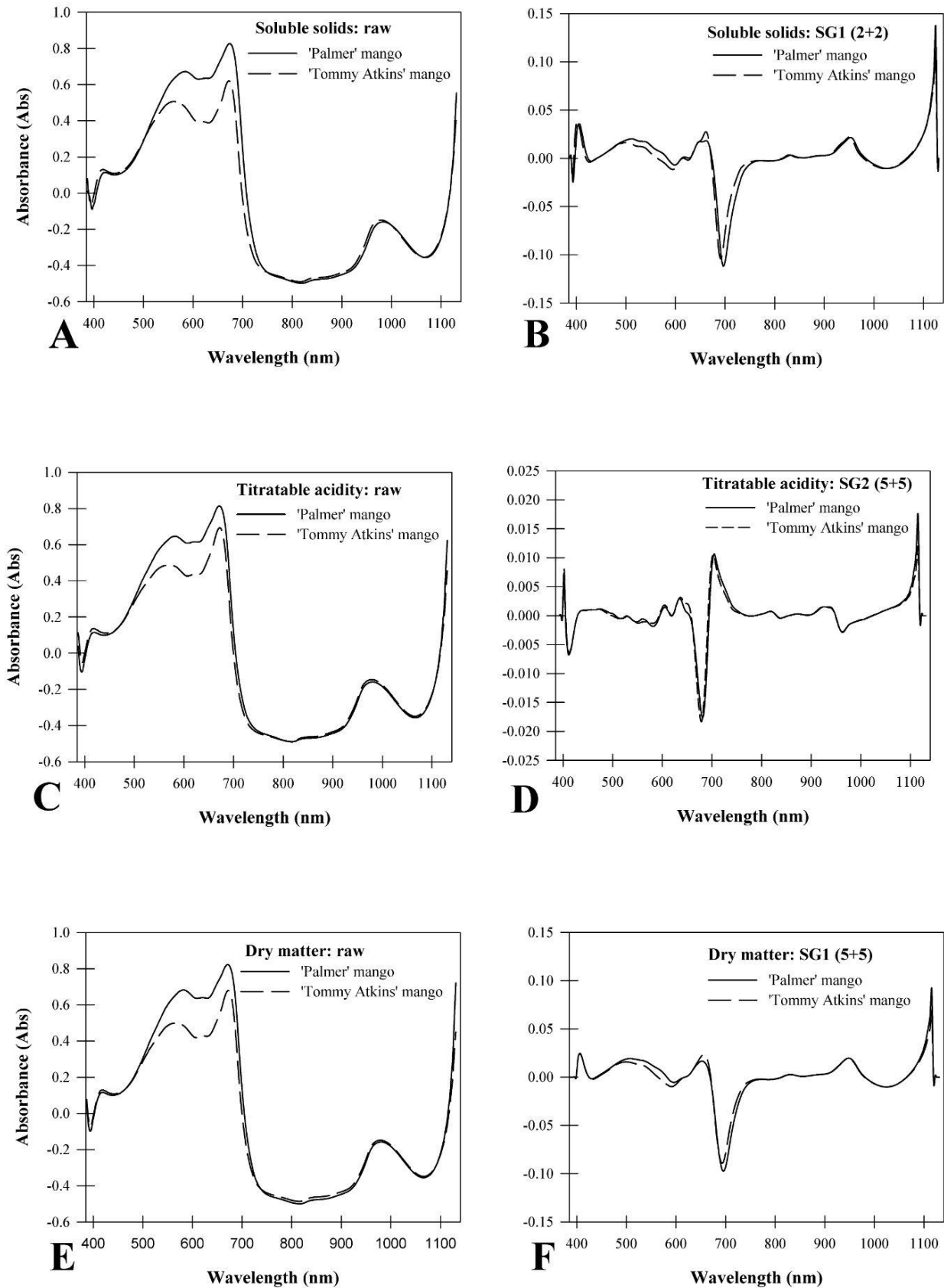
For 'Nam Dok Mai' mangoes, the SEP values determined ranged from 0.9 to 1.2% in the 700–1,100 nm range. This is similar to the low  $RMSE_p$  value found for 'Tommy Atkins' mangoes where 0.92% was obtained in the 950–1650 nm range by Marques et al. (2016).

### 3.4 Titratable acidity

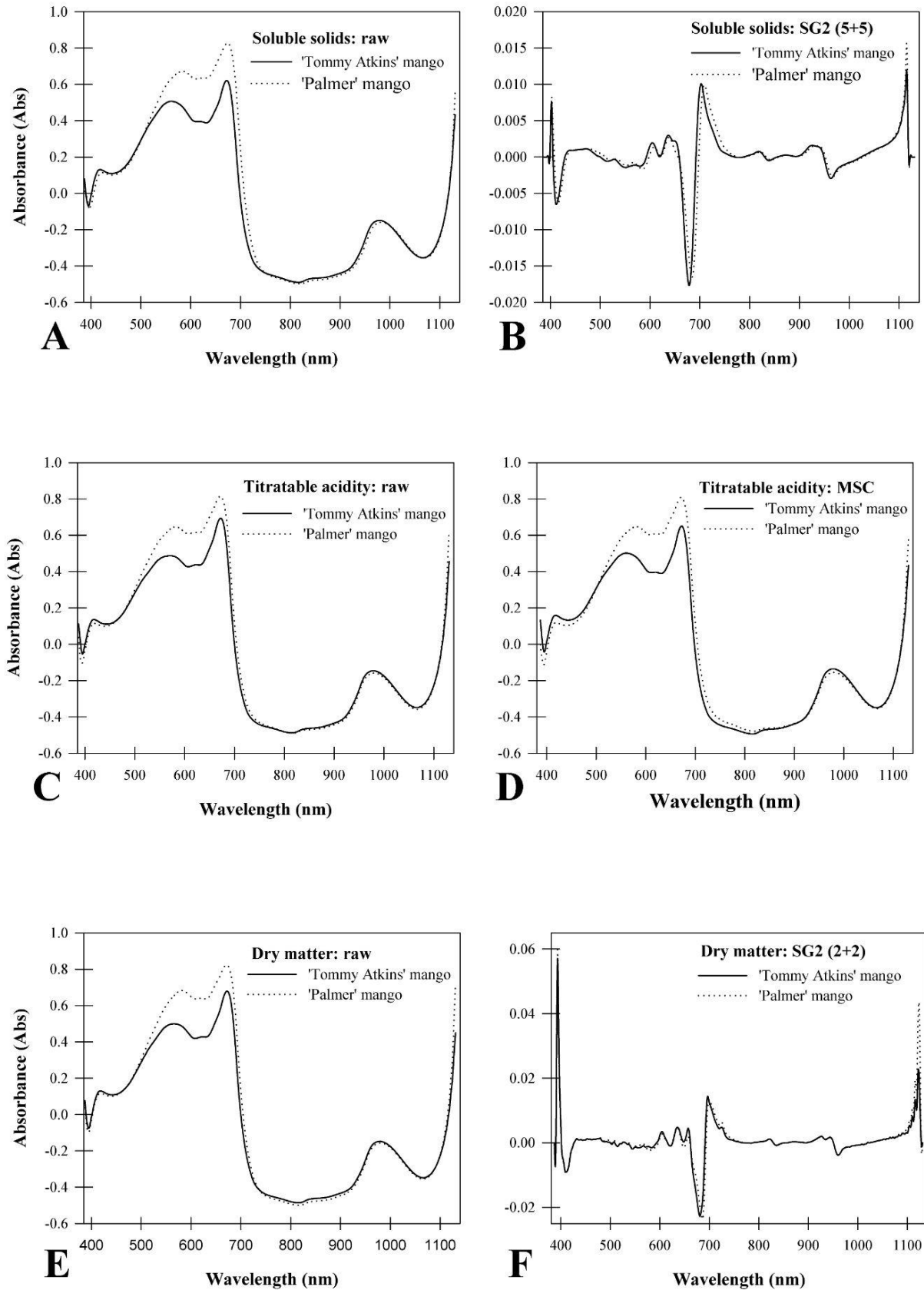
The SG2 pre-treatment with a 5+5 smoothing window using the 744–903 nm spectral window yielded prediction results for TA in 'Palmer' mangoes ( $R^2_p = 0.56$ ,  $RMSE_p = 0.14\%$ , and  $SDR_p = 1.48$ ) (Table 3), generating a calibration model with seven LVs,  $R^2_c = 0.64$ , and  $RMSE_{cv} = 0.13\%$ . In the case of 'Tommy Atkins' mangoes, the best TA prediction model ( $R^2_p = 0.59$ ,  $RMSE_p = 0.13\%$ , and  $SDR_p = 1.55$ ) occurred with MSC pre-processing employing five LVs, resulting in  $R^2_c = 0.63$  and  $RMSE_{cv} = 0.14\%$  in the 744–903 nm spectral window (Table 3). These models exhibit low accuracy (Nicolai et al., 2007), a common challenge faced while predicting acidity due to its typically low content in fruits, usually < 1%. In this study, the average TA was approximately 0.51% citric acid in the evaluated populations ranging from a minimum of 0.04% to a maximum of 1.12%. The difficulty in accurately predicting extreme acidity values is acknowledged by Nordey et al. (2017) who reported acidity as being influenced by ripening and environmental conditions, impacting the ratio of two main organic acids of mangoes (malic and citric acids).



**Figure 5.** (A) The principal component analysis results of 'Palmer' mangoes' dry matter in relative absorbance and (B) with the application of the Savitzky–Golay first derivative (5+5); (C) the principal component analysis results of 'Tommy Atkins' mangoes' dry matter in relative absorbance and (D) with the application of the Savitzky–Golay second derivative (2+2).



**Figure 6.** (A) Raw near-infrared spectra collected for the soluble solids of 'Palmer' mangoes and (B) with the application of the Savitzky–Golay first derivative (2+2); (C) raw near-infrared spectra collected for the titratable acidity of 'Palmer' mangoes and (D) with the application of the Savitzky–Golay second derivative (5+5); (E) raw near-infrared spectra collected for the dry matter of 'Palmer' mangoes and (F) with the application of the Savitzky–Golay first derivative (5+5).



**Figure 7.** (A) Raw near-infrared spectra collected for the soluble solids of 'Tommy Atkins' mangoes and (B) with the application of the Savitzky-Golay second derivative (5+5); (C) raw near-infrared spectra collected for the titratable acidity of 'Tommy Atkins' mangoes and (D) with the application of the multiplicative signal correction; (E) raw near-infrared spectra collected for the dry matter of 'Tommy Atkins' mangoes and (F) with the application of the Savitzky-Golay second derivative (5+5).



Marques et al. (2016) obtained models for ‘Tommy Atkins’ mangoes utilizing SNV pre-processing, reporting  $R^2_p = 0.50$  and  $RMSE_p = 0.17\%$ . Schmilovitchs et al. (2000) analyzing ‘Tommy Atkins’ mangoes within the TA range of 0.15–0.85% reported  $SEP = 0.22\%$  and  $R^2_p = 0.40$ .

### 3.5 Dry matter

In ‘Palmer’ mangoes, superior DM results were achieved with SG1 using a 5+5 smoothing window in the 699–981 nm spectral window: a calibration model with 11 LVs,  $R^2_c = 0.78$ , and  $RMSE_c = 0.88\%$  and a prediction model with  $R^2_p = 0.83$ ,

$RMSE_p = 0.95$ , and  $SDR_p = 2.44$  (Table 4). For ‘Tommy Atkins’ mangoes, optimal DM results were obtained with SG2 using a 2+2 smoothing window in the 699–981 nm spectral window: a calibration model with seven LVs,  $R^2_c = 0.87$ , and  $RMSE_c = 1.37\%$  and a prediction model with  $R^2_p = 0.79$ ,  $RMSE_p = 1.36$ , and  $SDR_p = 2.00$  (Table 4).

The DM models demonstrated superiority over SS and TA, particularly in terms of SDR, indicating enhanced accuracy in providing gross quantitative predictions (Nicolai et al., 2007). This could be attributed to the absorption of water and carbohydrates in the 960–990 nm range, as during maturation,

**Table 2.** Calibration and prediction results for soluble solids contents (%) in ‘Palmer’ and ‘Tommy Atkins’ mangoes.

cv.	Data treatment	Range, $\lambda$ (nm)	LV	Calibration model						Prediction model				
				$SD_c$	$R^2_c$	$R^2_{cv}$	$RMSE_c$	$RMSE_{cv}$	$^E SDR_c$	$SD_p$	$R^2_p$	$^A RMSE_p$	SEP	$^E SDR_p$
Palmer	<sup>B</sup> SG1 2+2	699–999	12	3.02	0.74	0.55	1.48	1.97	1.49	2.82	0.69	1.56	1.57	1.80
Palmer	<sup>C</sup> SG2 2+2	699–999	12	3.02	0.85	0.55	1.21	1.95	1.49	2.82	0.70	1.56	1.58	1.79
Palmer	<sup>D</sup> MSC	699–999	12	3.02	0.74	0.59	1.52	1.91	1.56	2.82	0.67	1.63	1.64	1.73
Tommy Atkins	SG2 5+5	699–999	7	4.61	0.89	0.86	1.46	1.73	2.66	4.41	0.72	2.43	2.38	1.85
Tommy Atkins	SG1 5+5	699–999	7	4.48	0.87	0.84	1.70	1.81	2.48	4.41	0.70	2.45	2.47	1.79
Tommy Atkins	MSC	699–999	7	4.63	0.87	0.83	1.65	1.89	2.45	4.41	0.69	2.46	2.49	1.77

$SD_c$ : standard deviation of calibration;  $R^2_c$ : coefficient of determination for calibration;  $R^2_{cv}$ : coefficient of determination for cross-validation;  $RMSE_c$ : square root of the mean error of calibration;  $RMSE_{cv}$ : square root of the mean error of cross-validation;  $SDR_c$ : standard deviation ratio of calibration;  $SD_p$ : standard deviation of prediction;  $R^2_p$ : coefficient of determination for prediction;  $RMSE_p$ : square root of the mean error of prediction; SEP: standard error of prediction;  $SDR_p$ : standard deviation ratio of prediction.  $^A RMSE_p$ : revised by bias;  $^B$ Savitzky–Golay first derivative (SG1);  $^C$ Savitzky–Golay second derivative (SG2);  $^D$ multiplicative signal correction (MSC);  $^E SDR_c = (SD_c/RMSE_c)$ ;  $^E SDR_p = (SD_p/RMSE_p)$ .

**Table 3.** Calibration and prediction results for titratable acidity (% citric acid) in ‘Palmer’ and ‘Tommy Atkins’ mangoes.

cv.	Data treatment	Range, $\lambda$ (nm)	LV	Calibration model						Prediction model				
				$SD_c$	$R^2_c$	$R^2_{cv}$	$RMSE_c$	$RMSE_{cv}$	$^E SDR_c$	$SD_p$	$R^2_p$	$^A RMSE_p$	SEP	$^E SDR_p$
Palmer	<sup>B</sup> SG2 5+5	744–903	7	0.19	0.64	0.51	0.11	0.13	1.44	0.195	0.56	0.14	0.13	1.48
Palmer	<sup>C</sup> SG2 2+2	744–903	4	0.19	0.54	0.40	0.12	0.14	1.32	0.195	0.57	0.14	0.13	1.47
Palmer	<sup>D</sup> MSC	744–903	6	0.18	0.60	0.48	0.11	0.13	1.44	0.195	0.51	0.15	0.14	1.41
Tommy Atkins	MSC	744–903	5	0.22	0.63	0.59	0.13	0.14	1.54	0.205	0.58	0.13	0.13	1.55
Tommy Atkins	SG2 5+5	744–903	6	0.21	0.77	0.65	0.10	0.12	1.70	0.205	0.56	0.15	0.14	1.50
Tommy Atkins	SG2 2+2	744–903	3	0.21	0.63	0.59	0.13	0.14	1.53	0.205	0.48	0.16	0.15	1.38

$SD_c$ : standard deviation of calibration;  $R^2_c$ : coefficient of determination for calibration;  $R^2_{cv}$ : coefficient of determination for cross-validation;  $RMSE_c$ : square root of the mean error of calibration;  $RMSE_{cv}$ : square root of the mean error of cross-validation;  $SDR_c$ : standard deviation ratio of calibration;  $SD_p$ : standard deviation of prediction;  $R^2_p$ : coefficient of determination for prediction;  $RMSE_p$ : square root of the mean error of prediction; SEP: standard error of prediction;  $SDR_p$ : standard deviation ratio of prediction.  $^A RMSE_p$ : revised by bias;  $^B$ Savitzky–Golay first derivative (SG1);  $^C$ Savitzky–Golay second derivative (SG2);  $^D$ multiplicative signal correction (MSC);  $^E SDR_c = (SD_c/RMSE_c)$ ;  $^E SDR_p = (SD_p/RMSE_p)$ .

**Table 4.** Calibration and prediction results for dry matter (%) in ‘Palmer’ and ‘Tommy Atkins’ mangoes.

cv.	Data treatment	Range $\lambda$ (nm)	LV	Calibration model						Prediction model				
				$SD_c$	$R^2_c$	$R^2_{cv}$	$RMSE_c$	$RMSE_{cv}$	$^D SDR_c$	$SD_p$	$R^2_p$	$^A RMSE_p$	SEP	$^E SDR_p$
Palmer	<sup>B</sup> SG1 5+5	699–981	11	1.90	0.78	0.75	0.88	0.96	1.98	2.23	0.83	0.95	0.91	2.44
Palmer	<sup>C</sup> SG2 5+5	699–981	8	1.90	0.75	0.69	0.94	1.04	1.82	2.23	0.81	1.01	0.98	2.27
Palmer	SG1 2+2	699–981	12	1.90	0.81	0.75	0.83	0.95	1.99	2.23	0.81	1.01	0.99	2.25
Tommy Atkins	SG2 2+2	699–981	7	3.79	0.87	0.80	1.37	1.73	2.19	3.024	0.79	1.54	1.51	2.00
Tommy Atkins	SG2 5+5	699–981	7	3.69	0.84	0.79	1.49	1.67	2.21	3.024	0.76	1.61	1.59	1.91
Tommy Atkins	SG1 2+2	699–981	6	3.75	0.81	0.78	1.62	1.77	2.12	3.024	0.69	1.95	1.87	1.62

$SD_c$ : standard deviation of calibration;  $R^2_c$ : coefficient of determination for calibration;  $R^2_{cv}$ : coefficient of determination for cross-validation;  $RMSE_c$ : square root of the mean error of calibration;  $RMSE_{cv}$ : square root of the mean error of cross-validation;  $SDR_c$ : standard deviation ratio of calibration;  $SD_p$ : standard deviation of prediction;  $R^2_p$ : coefficient of determination for prediction;  $RMSE_p$ : square root of the mean error of prediction; SEP: standard error of prediction;  $SDR_p$ : standard deviation ratio of prediction.  $^A RMSE_p$ : revised by bias;  $^B$ Savitzky–Golay first derivative (SG1);  $^C$ Savitzky–Golay second derivative (SG2);  $^D SDR_c = (SD_c/RMSE_c)$ ;  $^E SDR_p = (SD_p/RMSE_p)$ .

starch and DM accumulate and the moisture content gradually decreases (Santos Neto et al., 2017).

Mesocarp starch transforms into soluble sugars during mango ripening, with the total carbohydrate content (total starch and SS) measured by the DM serving as a suitable parameter connected to horticultural maturity and representing the final quality for consumption. Freitas et al. (2022) found a positive linear relationship between 'Palmer' and 'Tommy Atkins' mangoes' DM content at harvest and consumer satisfaction, suggesting that higher DM content results in increased overall satisfaction. Thus, mangoes' DM content serves as a valuable harvest index and guide for assessing the consumption quality of ripe fruits.

## 4 CONCLUSIONS

This study reaffirmed the considerable potential of NIR spectroscopy in assessing the internal quality of mangoes for precise harvesting decisions. NIR spectrometry proved effective in developing robust prediction models for DM content. However, the accuracy of TA prediction models for both 'Palmer' and 'Tommy Atkins' mangoes is relatively low. Nevertheless, technology holds promise for non-destructive, field-based mango quality predictions, contributing to informed decision-making.

## ACKNOWLEDGMENTS

We are immensely grateful to FAPEG for their crucial support in acquiring the F-750 portable NIR spectrometer. This contribution has been fundamental to the advancement of our research and the strengthening of science and innovation in our state. FAPEG's support has been essential for the development of new knowledge and technologies, positively impacting society. We are deeply thankful for this support and for FAPEG's commitment to scientific progress.

## REFERENCES

- Anderson, N. T., Walsh, K. B., Subedi, P. P., & Hayes, C. H. (2020). Achieving robustness across season, location and cultivar for a NIRS model for intact mango fruit dry matter content. *Postharvest Biology and Technology*, 168, 111202. <https://doi.org/10.1016/j.postharvbio.2020.111202>
- Association of Official Analytical Chemists (AOAC) (1997). *Official methods of analysis of the Association of Official Analytical Chemists International*. 16th ed. AOAC.
- Brito, A. A., Campos, F., Nascimento, A. R., Damiani, C., Silva, F. A., Teixeira, G. H. A., & Cunha Junior, L. C. (2022). Non-destructive determination of color, titratable acidity, and dry matter in intact tomatoes using a portable Vis-NIR spectrometer. *Journal of Food Composition and Analysis*, 107, 104288. <https://doi.org/10.1016/j.jfca.2021.104288>
- Fan, S., Wang, Q., Tian, X., Yang, G., Xia, Y., Li, J., & Huang, W. (2020). Non-destructive evaluation of soluble solids content of apples using a developed portable VIS/NIR device. *Biosystems Engineering*, 193, 138-148. <https://doi.org/10.1016/j.biosystemseng.2020.02.017>
- Freitas, S. T., Guimarães, I. T., Vilvert, J. C., Amaral, M. H. P., Brecht, J. K., & Marques, A. T. B. (2022). Mango dry matter content at harvest to achieve high consumer quality of different cultivars in different growing seasons. *Postharvest Biology and Technology*, 189, 111917. <https://doi.org/10.1016/j.postharvbio.2022.111917>
- Gianguzzi, G., Farina, V., Inglese, P. & Rodrigo, M. G. L. (2021). Effect of harvest date on mango (*Mangifera indica* L. cultivar Osteen) fruit's qualitative development, shelf life and consumer acceptance. *Agronomy*, 11(4), 811. <https://doi.org/10.3390/agronomy11040811>
- Golic, M., & Walsh, K. B. (2006). Robustness of calibration models based on near infrared spectroscopy for the in-line grading of stonefruit for total soluble solids content. *Analytica Chimica Acta*, 555(2), 286-291. <https://doi.org/10.1016/j.aca.2005.09.014>
- Lauricella, M., Emanuele, S., Calvaruso, G., Giuliano, M., & D'Anneo, A. (2017). Multifaceted health benefits of *Mangifera indica* L. (mango): the inestimable value of orchards recently planted in Sicilian rural areas. *Nutrients*, 9(5), 525. <https://doi.org/10.3390/nu9050525>
- Li, X., Wei, Y., Xu, J., Feng, X., Wu, F., Zhou, R., Jin, J., Xu, K., Yu, X., & He, Y. (2018). SSC and pH for sweet assessment and maturity classification of harvested cherry fruit based on NIR hyperspectral imaging technology. *Postharvest Biology and Technology*, 143(4), 112-118. <https://doi.org/10.1016/j.postharvbio.2018.05.003>
- Lobo, M. G., & Sidhu, J. S. (2017). Biology, postharvest physiology, and biochemistry of mango. In M. Siddiq, J. K. Brecht & J. S. Sidhu (Eds.), *Handbook of mango fruit: production, postharvest science, processing technology and nutrition* (pp. 37-59). John Wiley & Sons.
- Maldonado-Celis, M. E., Yahia, E. M., Bedoya, R., Landázuri, P., Loango, N., Aguillón, J., Restrepo, B., & Guerrero Ospina, J. C. (2019). Chemical composition of mango (*Mangifera indica* L.) fruit: nutritional and phytochemical compounds. *Front Plant Science*, 10, 1073. <https://doi.org/10.3389/fpls.2019.01073>
- Marques, E. J. N., 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(Part B), 1207-1214. <https://doi.org/10.1016/j.foodchem.2015.11.080>
- Ministério da Agricultura, Pecuária e Abastecimento (MAPA) (2023). *AGROSTAT - Estatísticas de Comércio Exterior de Agronegócio Brasileiro*. Retrieved from <https://indicadores.agricultura.gov.br/agrostat/index.htm>
- Mishra, P., & Woltering, E. (2023). Semi-supervised robust models for predicting dry matter in mango fruit with near-infrared spectroscopy. *Postharvest Biology and Technology*, 202, 112335. <https://doi.org/10.1016/j.postharvbio.2023.112335>
- Munawar, A. A., Zulfahrizal, Meilina, H., & Pawelzik, E. (2022). Near infrared spectroscopy as a fast and non-destructive technique for total acidity prediction of intact mango: Comparison among regression approaches. *Computers and Electronics in Agriculture*, 193(2), 106657. <https://doi.org/10.1016/j.compag.2021.106657>
- Nicolai, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saey, 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(2), 99-118. <https://doi.org/10.1016/j.postharvbio.2007.06.024>
- 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. <https://doi.org/10.1016/j.scienta.2016.12.023>
- Santos Neto, J. P., Assis, M. W. D., Casagrande, I. P., Cunha Junior, L. C., & Teixeira, G. H. A. (2017). Determination of 'Palmer' mango maturity indices using portable near infrared (VIS-NIR) spectrometer. *Postharvest Biology and Technology*, 130, 75-80. <https://doi.org/10.1016/j.postharvbio.2017.03.009>

- Schmilovitchs, Z., Mizrach, A., Hoffman, A., Egozi, H., & Fuchs, Y. (2000). Determination of mango physiological indices by near-infrared spectrometry. *Postharvest Biology and Technology*, 19(3), 245-252. [https://doi.org/10.1016/S0925-5214\(00\)00102-2](https://doi.org/10.1016/S0925-5214(00)00102-2)
- Shah, S. S. A., Zeb, A., Qureshi, W. S., Malik, A. U., Tiwama, M., Walsh, K., Amin, M., Alsamary, W., & Alanazi, E. (2021). Mango maturity classification instead of maturity index estimation: A new approach towards handheld NIR spectroscopy. *Infrared Physics & Technology*, 115, 103639. <https://doi.org/10.1016/j.infrared.2021.103639>
- Subedi, P. P., & Walsh, K. B. (2011). Assessment of sugar and starch in intact banana and mango fruit by SWNIR spectroscopy. *Postharvest Biology and Technology*, 62(3), 238-245. <https://doi.org/10.1016/j.postharvbio.2011.06.014>
- Subedi, P. P., Walsh, K. B., & Hopkins, S. W. (2012). Assessment of titratable acidity in fruit using short wave near infrared spectroscopy. Part A: Establishing a detection limit based on model solutions. *Journal of Near Infrared Spectroscopy*, 20(4), 449-457. <https://doi.org/10.1255/jnirs.1010>
- Subedi, P. 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(3), 326-334. <https://doi.org/10.1016/j.postharvbio.2006.09.012>
- Toscano, G., Rinnan, Å., Pizzi, A., & Mancini, M. (2017). The use of near-infrared (NIR) spectroscopy and principal component analysis (PCA) to discriminate bark and wood of the most common species of the pellet sector. *Energy Fuels*, 31(3), 2814-2821. <https://doi.org/10.1021/acs.energyfuels.6b02421>
- Yahia, E. M. (2019). Mango (*Mangifera indica* L.). In E. M. Yahia (Ed.), *Postharvest biology and technology of tropical and subtropical fruits: cocona to mango* (pp. 492-565). Woodhead Publishing Limited.
- Yu, Y., & Yao, M. (2022). A portable NIR system for nondestructive assessment of SSC and firmness of Nanguo pears. *LWT*, 167, 113809. <https://doi.org/10.1016/j.lwt.2022.113809>
- Zhang, W., Zhu, G., & Zhu, G. (2022). The imitation and creation of a mango flavor. *Food Science and Technology*, 42, e34622. <https://doi.org/10.1590/fst.34622>