Applied Food Research 5 (2025) 101553

Contents lists available at ScienceDirect

Applied Food Research

journal homepage: www.elsevier.com/locate/afres

Near infrared spectroscopy for sustainable non-destructive classification and quantitative prediction of quality traits of seed size, pest damage, and quality traits in lychee (*Litchi chinensis*)

Saowaluk Rungchang [a], Wouter Saeys [b], Chayanid Sringarm [c], Sonthaya Numthuam [d], Suphannika Intanon [d], Sila Kittiwachana [e], Sudarat Jiamyangyuen [a,*]

[a] Division of Food Science and Technology, Faculty of Agro-Industry, Chiang Mai University, Chiang Mai 50100, Thailand
[b] KU Leuven Department of Biosystems, MeBioS Unit, Kasteelpark Arenberg 30, 3001 Leuven, Belgium
[c] Department of Agro-Industrial, Food, and Environmental Technology, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok (KMUTNB), Bangkok, 10800, Thailand
[d] Faculty of Agriculture Natural Resources and Environment, Naresuan University, Phitsanulok, 65000, Thailand
[e] Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai, 50200, Thailand

ARTICLE INFO

Keywords:
Fruit physiology
Postharvest grading
Quality assessment
Chemometrics
NIR spectroscopy

ABSTRACT

Lychees are highly valued in export markets if they have an aborted seed and are free from fruit borer damage. However, the traditional grading method relies on manual assessment by laborers, which is experience-dependent, time-consuming, and potentially destructive to the fruit. The objective of this study was to evaluate NIR spectroscopy for classifying premium lychee fruits based on seed size and the absence of fruit borer damage, as well as for quantifying the soluble solids content (SSC) and titratable acidity (TA). NIR spectra in the 12,000–4000 cm^{-1} range were acquired in reflectance mode at three different positions: stem end, cheek side, and tip. The classification models effectively distinguished seed size at tip side, achieving 90.58 % accuracy and fruit borer damage at stem-end, resulting in 85.71 % and 90.91 % correct classification for normal and aborted seeds, respectively. Additionally, SSC and TA content were quantitatively assessed from the NIR spectra with R^2 values > 0.90. These results could pave the way for efficient quality control of lychees. This research advances innovative sensing technologies for agriculture while enabling sustainable production and resource-efficient food systems.

1. Introduction

Lychee (*Litchi chinensis* Sonn.) is a tropical fruit tree of the Sapindaceae family, a native to China. It is widely cultivated in subtropical regions such as the northern part of Thailand, southern China and northern Vietnam, and commonly consumed in many countries. A variety of taste, smell, and appearance including fruit shape, and seed characteristics of lychee fruits are available as a result of cultivar development (Raz. et al., 2022). In Thailand, during the peak season from April to June 2023, the export value of fresh and frozen lychees was approximately 1.6 million Thai Baht for an export volume of around 8.3 thousand kilograms. Satisfaction factors in purchase and consumption by consumers are related to fruit appearance and its internal quality. The external appearance includes size, shape, color, gloss, and the presence or absence of defects or decay. Customer acceptance of the internal quality depends on biochemical parameters such as a soluble solids content between 16 and 18 °Bx and a total acidity (TA) from 0.3 to 0.4 % (Bhushan et al., 2015). The seed size of lychee has become an important internal characteristic to meet consumer preferences for premium quality lychees. Seedlessness or minimal internal seed presence are demanded by the high-end and export markets. Fruit with an aborted seed typically have a higher proportion of edible flesh within the fruit, underscoring the importance of this seed characteristic as a key parameter for fruit quality (Chu et al., 2015; Zhang et al., 2018; Matsuda & Higuchi., 2019; Xie et al., 2019a). Furthermore, the absence of damage from pests and worms is also crucial to be graded as premium quality. Lychee fruit borer (*Conopomorpha sinensis* Bradley) is one of the most destructive pests that causes extensive loss to farmers. This pest typically infests lychee fruit during their reproductive growth phase,

* Corresponding author.
E-mail address: sudarat.j@cmu.ac.th (S. Jiamyangyuen).

https://doi.org/10.1016/j.afres.2025.101553
Received 1 August 2025; Received in revised form 17 November 2025; Accepted 24 November 2025
Available online 25 November 2025
2772-5022/© 2025 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Abbreviations

NIR	Near-Infrared
SSC	soluble solids content
TA	titratable acidity
NS w/o	normal seed without fruit borer
NS w	normal seed with fruit borer
AS w/o	aborted seed without fruit borer
AS w	aborted seed with fruit borer
PLS	Partial Least Squares
PLS-DA	Partial Least Squares–Discriminant Analysis
PLSR	Partial Least Squares Regression
iPLS	interval Partial Least Squares Regression

between March and May, causing initial damage to the nut before spreading to the flesh. Premium-grade lychees are graded through visual inspection by skilled workers, but the seed size and internal deterioration from fruit borer infestation are hard to detect through visual inspection.

Near-infrared (NIR) spectroscopy has been successfully used for quantitative analysis of fruit and vegetables (Nicolaï et al., 2014). It has for example been applied to determine the soluble solids in mango and cantaloupe and the sugar content in intact peaches and satsuma mandarin Saranwong et al. (2004); Masithoh et al. (2016); Nascimento et al. (2016); Ong et al. (2022). Moreover, it has been applied to detect microstructural properties such as the granulation in "Sai Num Pung" tangerine fruits (Theanjumpol et al., 2019) and puffiness in Satsuma mandarin Sun et al. (2023). Various classification methods have been employed for classifying food and agricultural products based on their NIR spectra, including partial least squares-discriminant analysis (PLS-DA), linear discriminant analysis, quadratic discriminant analysis (Saeys et al., 2019; Theanjumpol et al., 2019)

Based on the higher mention state of the art, it is hypothesized that NIRS could provide a more systematic, non-destructive approach for quality grading of lychee fruit. However, to assess the internal quality of intact lychee fruits, sufficient NIR penetration beyond the peel and into the pulp is essential. Given that penetration depth varies with wavelength and tissue properties, NIR spectra were measured at multiple positions on the fruit (stem end, cheek side, and tip), based on the hypothesis that anatomical differences at these sites might influence light scattering and penetration. These measurements aimed to identify the optimal site for detecting internal attributes such as seed size and fruit borer damage. Greensill & Walsh, (2000) reported that the intensity of the detected light in fruit decreases exponentially with distance from the source, following the Beer-Lambert law. As the attenuation in the tissue depends on its wavelength-dependent bulk optical properties, this depth also varies with wavelength (Lu et al., 2020). Lammertyn et al. (2000) showed that the penetration depth in 'Jonagold' apple fruit tissue is wavelength-dependent, reaching up to 4 mm in the 11,111–14,285 cm^{-1} range and between 2 and 3 mm in the 5263–11,111 cm^{-1} range. More recently, Wedding et al. (2024) studied the wavelength dependence of the penetration depth of NIR radiation in 'Hass' avocado fruit in reflectance mode. They found the maximum penetration depth of approximately 4.8 mm in the 11,625 to 11,987 cm^{-1} region and a lower penetration depth of 3.3 mm around 9311 cm^{-1}.

Given the increasing consumer demand for high-quality, seedless lychees in international markets, effective and rapid grading systems are crucial to maintain export standards and minimize postharvest losses. As Thailand's lychee export sector plays a significant economic role during the peak season, the development of non-destructive technologies for quality classification can directly support value chain efficiency and enhance competitiveness in the global fruit market. Therefore, the objective of this study was to investigate the potential of NIR spectroscopy as a non-destructive method for the classification of lychee premium fruits according to seed size and fruit borer damage, and for quantification of soluble solids content (SSC) and titratable acidity (TA) in lychee pulp. This proposed technique could assist premium grade lychee classification through a more standardized approach beneficial for the lychee export industry.

2. Materials and methods

2.1. Plant materials and sample preparation

Fresh lychee fruit (*Litchi chinensis* Sonn. cv. Pa Chid) were obtained from commercial orchards in Phetchabun Province, Thailand. Samples were collected during the peak maturation season (May-June) during two harvesting years (2023 and 2024). At the orchards, the harvested fruit were immediately pretreated by hydro-cooling, and then refrigerated at 4–10 °C for 12 h to delay the growth of microorganisms and to facilitate worm removal. After the chilling, skilled orchard workers visually inspected and graded the lychee fruit based on their seed size and the presence or absence of worm damage. Physical quality sorting was performed by gently squeezing the fruit's tip sideways to differentiate between shriveled and normal seeds, while worm damage was assessed by inspecting the stem area. A total of 932 lychee fruits were categorized into two groups based on seed size. These groups were further separated based on the presence or absence of worm damage to obtain the following classes: normal seed without fruit borer, normal seed with fruit borer, aborted seed without fruit borer, and aborted seed with fruit borer, as illustrated in Fig. 1. The sample weights, seed sizes and physical characteristics of the lychee classes used in this study are summarized in Table 1.

2.2. Acquired NIR spectra

All lychee samples were sealed in plastic bags to prevent desiccation immediately after harvest and stored at 4–10 °C with non-contact ice in a box during transport to the Laboratory at Naresuan University, Phitsanulok province, Thailand (117 km from the fruit collection site). Upon arrival in the laboratory (on the same day) the fruits were kept in an incubator at 25 °C for 2 h prior to the NIR measurements. The instrument was equipped with a 1.0 m fiber optic diffuse reflectance probe and a thermoelectrically cooled InGaAs detector. The fiber optic cable collected light scattered in various directions and guided it to the spectrometer. The NIR spectra were recorded in reflectance mode using a Matrix-F Fourier transform NIR spectrometer (Bruker, Germany) in the 12,000–4000 cm^{-1} range at a spectral resolution of 16 cm^{-1} in a temperature controlled room at 25 °C. NIR spectra were acquired from 3 positions on each fruit: (A) stem end, (B) cheek side, and (C) tip (Fig. 2). Once the NIR spectra at these three locations had been measured, the fruit was peeled to measure the actual seed size using a digital vernier caliper and to evaluate the presence of damage by fruit borer.

2.3. Analysis of physico-chemical parameters

The soluble solids content (SSC) and titratable acidity (TA) of each individual lychee fruit were measured as indicators for their flavour quality. SSC was recorded using a digital refractometer (Mettler, Model PR-101, Switzerland) on juice extracted with a manual citrus press. TA was measured based on the AOAC Official Method 942.15 and expressed as mass percentage of malic acid. The analyses of SSC and TA for each sample were performed in triplicate, which allowed to calculate the standard error (SE). Each single lychee fruit, including three different parts: the stem-end (A), the cheek side (B), and the tip side (C) was measured separately for Soluble Solids Content (SSC) and Titratable Acidity (TA).

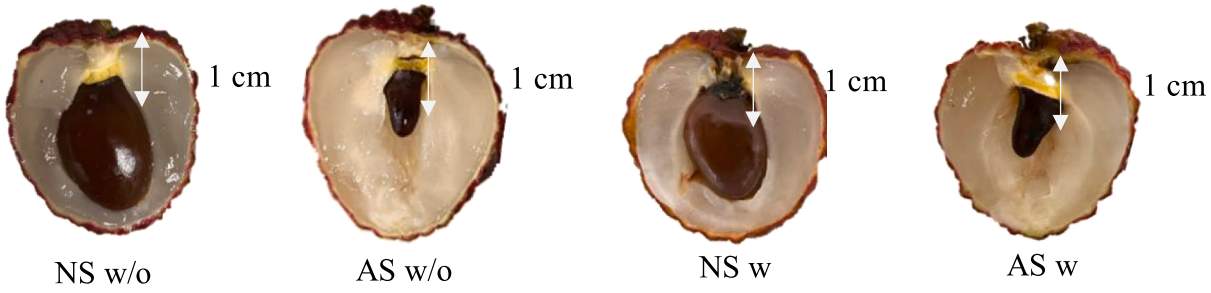


Fig. 1. Illustration of the 4 lychee classes graded according to seed size and damage caused by fruit borer: Class normal seed without fruit borer (NS w/o); Class aborted seed without fruit borer (AS w/o); Class normal seed with fruit borer (NS w); Class aborted seed with fruit borer (AS w).

Table 1
Sample weights, seed sizes and physical characteristics of the lychee classes.

	Number of samples	Description	Weight (g)	Seed size (width in mm)
Normal seed (NS)	233	without fruit borer (w/o)	31.03 ± 0.69	16.83 ± 0.80
	233	with fruit borer (w)	29.17 ± 0.11	15.51 ± 0.10
Aborted seed (AS)	233	without fruit borer (w/o)	31.19 ± 0.18	0.82 ± 0.10
	233	with fruit borer (w)	26.08 ± 0.75	0.81 ± 0.05

2.4. Chemometric analysis

First, classification models were built to grade the lychee samples into 4 classes, according to their seed size and worm damage based on the acquired NIR spectra. Next, quantification models were built to predict the SSC and TA from the acquired spectra.

2.4.1. Classification model for lychee grading

A hierarchical classifier with two levels was built. A first classification model was built to discriminate lychee fruit according to their seed size. Next, the fruit within each class were further separated according to the presence or absence of fruit borer damage. The classification models were developed using PLS-DA, which is an extension of partial least squares regression (PLSR) designed to identify differences among sample classes (Saeys et al., 2019). The 466 normal and 466 aborted fruit samples (932 in total) were each randomly divided into 312 calibration and 154 test samples. Consequently, the balanced calibration set comprised 624 fruits, while the balanced validation set contained 308 fruits. The number of latent variables (LVs) was selected as the one that gave the lowest RMSE value in a venetian blinds cross-validation (CV) with 10 splits and 1 sample per blind. Variable selection was then performed using interval partial least squares regression (iPLS) to identify the most informative spectral regions (Nørgaard et al., 2000). Specifically, a forward iPLS strategy was employed, initially dividing the spectra into 20 broad intervals. In each iteration, the most relevant intervals were progressively refined by splitting them into smaller, equal-width sub-intervals to improve model performance. Once the classification model had been trained on the 624 calibration samples, it was tested on the remaining 308 fruit from the test set to evaluate its performance in terms of the following metrics: (i) accuracy, representing the ratio of correctly classified samples to the total number of samples; (ii) sensitivity, indicating the proportion of correctly classified positive samples; and (iii) specificity, denoting the proportion of correctly classified negative samples (Melado-Herreros et al., 2021). The F1-score was computed as the harmonic mean of precision and sensitivity, with values ranging from 0 to 1. Higher F1-scores indicate superior classification performance (Melado-Herreros et al., 2021; Dong et al., 2024). All calculations were carried out in Matlab R2020a (The Mathworks, Natick, USA). Multivariate data analysis was conducted using the PLS Toolbox (Eigenvector Research, Wenatchee, WA, USA).

2.4.2. Regression models for SSC and TA

PLSR was employed to predict the soluble solid content (SSC) and Titratable Acidity (TA) values of the lychee fruit from the acquired NIR spectra. This analysis was limited to the 466 lychee samples without worm damage (class 1 and 2). The reference values for these samples are summarized in Table 8. All spectra were analyzed using the PLSR model with wavelength selection based on the iPLS method, which reduced dimensionality by removing irrelevant and redundant wavelengths (Dong et al., 2025). The optimal number of latent variables was determined based on 10-fold cross-validation, selecting the number that resulted in the minimum RMSECV for each model. (Porep, Kammerer & Carle, 2015; Sringarm et al., 2022). Once the model with the selected complexity had been trained on the full training set ($n = 311$), its prediction performance was evaluated on the test set ($n = 155$) and quantified in terms of the root mean squared error of prediction (RMSEP), the relative root mean square error of prediction (RRMSEP), and the coefficient of determination in prediction (R_{p2}) (Wang & Liu, 2024).

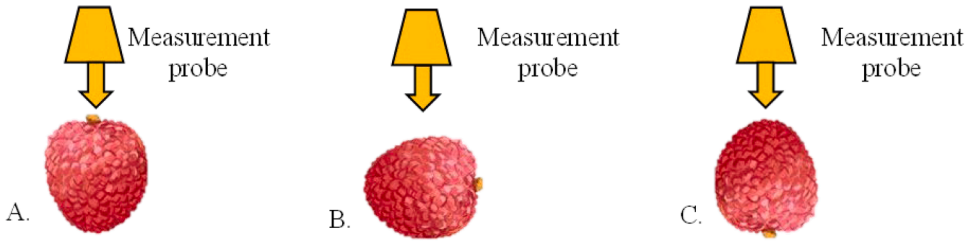


Fig. 2. Schematic illustration of the NIR measurements at 3 locations on the lychee fruit: stem end (A), cheek side (B), and tip (C) position.

3. Results and discussion

3.1. NIR spectra of Lychee fruit

The average Log(1/R) transformed diffuse reflectance spectra for the different classes and measurement locations on the lychee fruit are illustrated in Fig. 3. The peaks around 5300 cm^{-1} and 6900–7000 cm^{-1} may respectively be attributed to the combination band and first overtone of the stretching vibrations of O—H in water (Sringarm et al., 2024a; Bázár et al., 2016; Manley, 2014). The peaks around 4800, 4348 and 5586 cm^{-1} have been attributed to the stretching vibration of C—H, which has been reported to be indicative for the soluble solids content (Reddy & Yadav, 2018; Magwaza & Opara, 2015).

3.2. Lychee grading according to seed size and fruit borer damage using NIRS

Table 2 presents the confusion matrices for the test set obtained with the PLS-DA models distinguishing between fruit with normal and aborted seeds based on NIR spectra acquired at the stem end, cheek-side, and tip of the lychee fruit. The best classification performance, with balanced accuracy (90.58 %) was obtained based on the spectra acquired from the tip side of the lychee fruit. As shown in Table 3, these models obtained sensitivity scores, specificity scores, and F1-score of 0.86, 0.95 and 0.90, respectively, for tip position. This could be attributed to the physiological differences between fruit with normal and aborted seeds, as fruit with an aborted seed have a more extensive empty cavity and air pores in the lychee flesh, leading to increased light scattering (Kumar et al., 2017; Reddy & Yadav, 2018). However, the classification results obtained from spectra acquired at the stem end and cheek positions were still quite good with balanced accuracies of 76.94 % and 79.22 %, respectively. Fig. 4(a) presents the ROC curve for classifying lychee fruit seeds as normal (NS) or aborted (AS). The area under the ROC curve (AUC) was 0.967, demonstrating an excellent level of discrimination ability and confirming that the developed model provides highly accurate classification results. Fig. 4(b) shows the effectiveness of forward iPLS in selecting informative spectral intervals for classification. The most relevant spectral intervals for differentiating between normal and aborted lychee seeds are primarily concentrated in the region of 4800–4900 cm^{-1}, which is associated with C—H combination bands related to sugars. The region around 6000 cm^{-1} corresponds to the first overtone of C—H stretching vibrations in CH_3 and CH_2 functional groups (Xie et al., 2009; Goodarzi et al., 2015). Additionally, the regions near 7500 and 9500 cm^{-1} are attributed to the second overtone of O—H stretching vibrations, commonly related to hydroxyl groups (Li et al., 2017). These spectral regions contain critical information for distinguishing between the two seed types. The observed differences may be due to variations in sugar content, with lychee fruits containing aborted seeds potentially having higher sugar accumulation than those with normal seeds (Kumar et al., 2017).

After classifying the lychee fruit based on the seed size, the samples within each class (normal and aborted seed) were further classified according to the presence of fruit borer damage: Class NS without fruit borer (n = 233) comprises lychees without damage from worms or insects, while Class NS with fruit borer (n = 233) includes lychees with damage from worms or insects. The confusion matrices for fruit borer classification in Table 4 show differences in classification performance between samples with normal seeds with and without fruit borer damage based on the NIR spectra acquired at different positions on the lychee fruit. The best classification performance, with balanced accuracy (85.71 % accuracy) was obtained using the spectra from the stem end position of the lychee. The best model using spectra from the stem end position combines a sensitivity score of 0.93 for the normal seed without fruit borer class and 0.90 for the normal seed with fruit borer class. The F1-scores were 0.92 and 0.91 for the normal seed without fruit borer and with fruit borer classes (Table 5). This may be due to the fact that lychee fruit borers mainly cause damage close to the stem end. The presence of fruit borers can lead to changes in the texture, making it softer or mushier in areas where the larvae have fed, which affects the overall firmness of the fruit. Additionally, the chemical composition of the fruit may change due to the activity of the fruit borers, leading to a decrease in sugar content and altering the taste profile. As a result, NIR light can penetrate deeper into the fruit (de Oliveira et al., 2014), allowing it to capture information on both the damaged areas and the size of the flesh, particularly from the stem end to the interior. Although the lychee pericarp is relatively thick (typically 1.5–2.5 mm), the successful discrimination between lychees according to their seed size and the presence of internal damage indicates that the light penetration beyond the peel was sufficient to capture these internal differences. This is in line with the reports that NIR light in the 830–1100 nm range can penetrate up to 4–5 mm inside a fruit (Lambertyn et al., 2000). Fig. 5(a) presents the ROC curve for classifying lychee fruits with normal seeds either without or with worm infestation. The AUC was 0.962, indicating a very high level of discrimination ability and confirming the strong predictive performance of the developed model. Fig. 5(b). illustrates the results of forward interval partial least squares (iPLS), which was applied to select the most informative spectral intervals for the classification between normal seed without fruit borer and with fruit borer

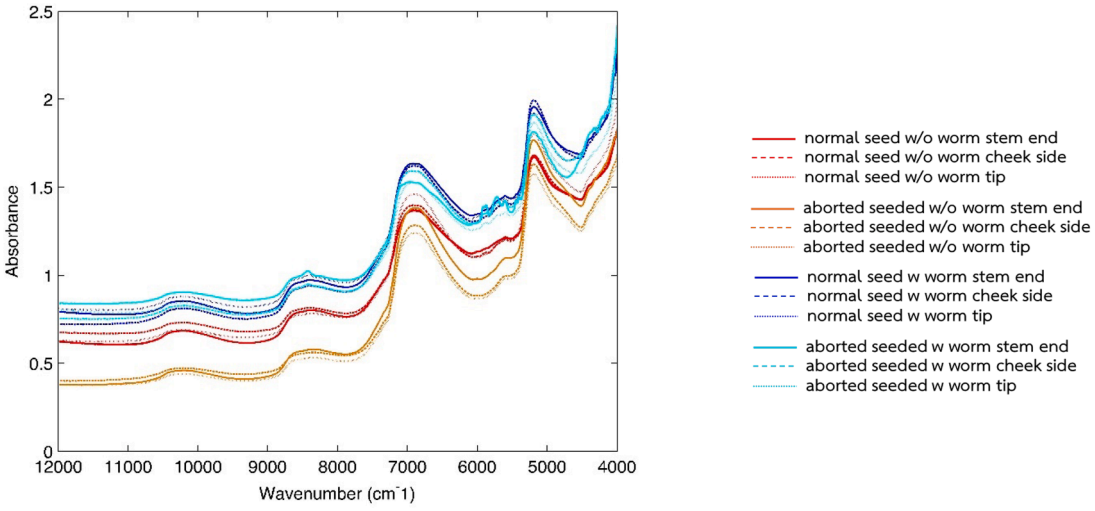


Fig. 3. Average near infrared (NIR) spectra of lychee fruit from the different quality classes, measured at the stem end (A), cheek side (B), and tip position (C).

Table 2
Confusion matrices for the test set obtained with the PLS-DA models based on variables selected using iPLS classifying lychee fruit as normal seed (NS) ($n = 154$) or aborted-seed (AS) ($n = 154$).

	A. Stem end		B. Cheek side		C. Tip	
LVs	6		9		8	
	Actual		Actual		Actual	
	Class NS	Class AS	Class NS	Class AS	Class NS	Class AS
Predicted as Class NS	103	20	116	26	133	8
Predicted as Class AS	51	134	37	128	21	146
Predicted as unassigned	0	0	0	0	0	0

Table 3
Performance metrics for the PLS-DA models based on variables selected using iPLS classifying lychee fruit in the classes normal seed (NS) and aborted seed (AS) according to the NIR spectra acquired at different positions on the lychee fruit.

	A. Spectra stem end			B. Spectra at cheek side			C. Spectra at tip		
	Sensitivity	Specificity	F1	Sensitivity	Specificity	F1	Sensitivity	Specificity	F1
Class NS	0.67	0.87	0.74	0.75	0.83	0.78	0.86	0.95	0.90
Class AS	0.87	0.67	0.79	0.83	0.75	0.80	0.95	0.86	0.90

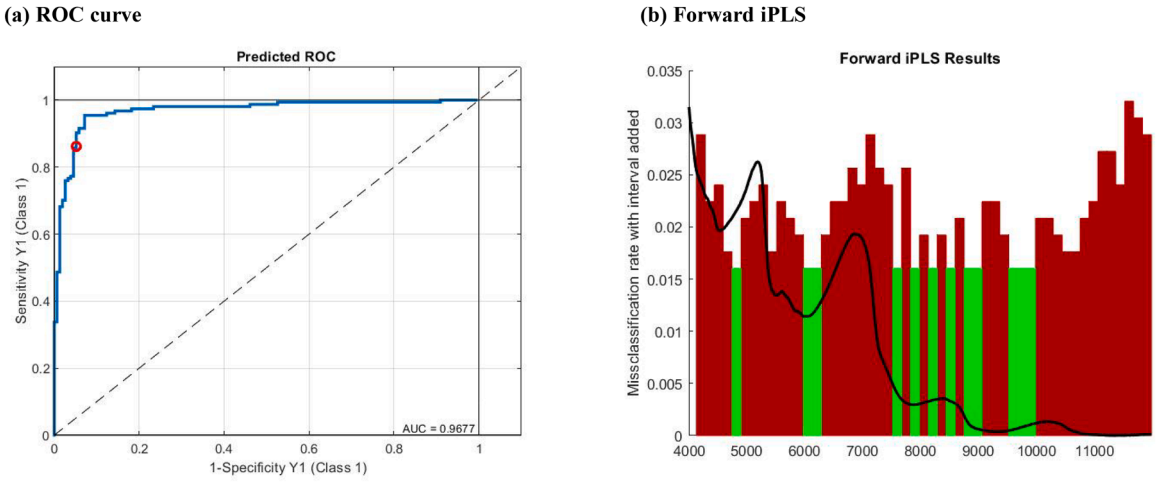


Fig. 4. (a) ROC curve and (b) Forward iPLS for classifying lychee fruit as normal seed (NS) and aborted-seed (AS) using spectra acquired at the tip.

Table 4
Confusion matrices for the PLS-DA models based on variables selected using iPLS classifying lychee fruit with normal seed (NS) as without fruit borer damage (NS w/o, $n = 77$) and with fruit borer damage (NS w, $n = 77$) based on the NIR spectra acquired at different locations on the fruit.

	A. Stem end		B. Cheek side		C. Tip	
LVs	6		4		5	
	Actual		Actual		Actual	
	Class NS w/o	Class NS w	Class NS w/o	Class NS w	Class NS w/o	Class NS w
Predicted as Class NS w/o	62	7	72	8	64	10
Predicted as Class NS w	15	70	5	69	13	67
Predicted as unassigned	0	0	0	0	0	0

classes. The most relevant spectral regions for characterizing chemical components in lychee fruit were identified around 4500 cm^{-1}, corresponding to the first C—H combination bands typically associated with sugars. The region near 6000 cm^{-1} relates to the first overtone of C—H stretching vibrations in CH_3 and CH_2 functional groups. Additionally, spectral features around 8700 cm^{-1} are attributed to the second overtone of C—H stretching, while the region between 10,500–11,200 cm^{-1} corresponds to the third overtone of CH_2 stretching vibrations (Kawano et al., 1992; Theanjumpol et al., 2019). Moreover, the region from 6900 to 7100 cm^{-1} is associated with the first overtone of the O—H stretching vibration in water (Xie et al., 2009). These spectral variations may be influenced by the activity of fruit borers, which can induce biochemical

Table 5
Performance metrics of the PLS-DA model in the prediction dataset based on variables selected using iPLS for normal seeded without fruit borer (NS w/o) and normal seeded with fruit borer (NS w).

	A. Spectra stem end			B. Spectra at cheek side			C. Spectra at tip			
	Sensitivity	Specificity	F1	Sensitivity	Specificity	F1	Sensitivity		Specificity	F1
Class NS w/o	0.93	0.90	0.92	0.81	0.91	0.85	0.83	0.87		0.85
Class NS w	0.90	0.93	0.91	0.91	0.81	0.86	0.87	0.83		0.85

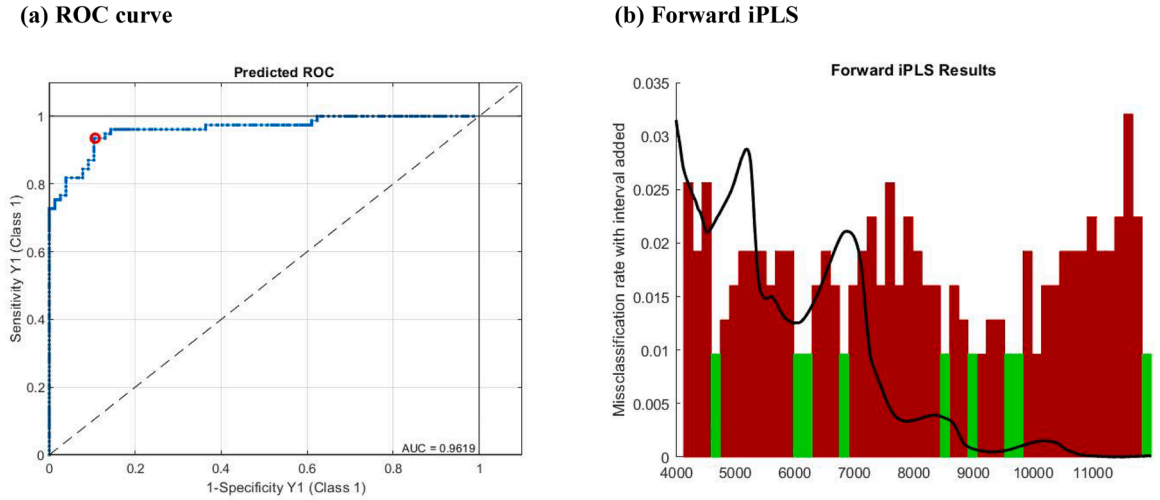


Fig. 5. (a) ROC curve and (b) Forward iPLS for classifying lychee fruit as normal seed without and with worm using spectra acquired at the stem end.

changes within the fruit. Such damage may lead to a reduction in sugar content and changes in moisture levels, thereby affecting the overall chemical composition of the fruit.

The classification results for fruit borer damage in lychee fruit with aborted seeds are summarized in Table 6. These models were built using spectra from different positions, covering Class AS without fruit borer (n = 233), including lychees without damage from worms or insects, and AS with fruit borer (n = 233), including lychees with damage from worms or insects. The best results were obtained from the stem-end position, compared to the cheek side and tip positions, with an average accuracy of 90.91 % across all classes and minimal sample misclassification. The corresponding performance metrics are summarized in Table 7. The highest sensitivity in fruit borer detection was obtained for the lychee fruit with aborted seeds (class AS), with specificity and F1-scores greater than 0.90, making them suitable for monitoring purposes and quality control in the lychee industry, ensuring consumers receive high-quality fruit. The ROC curve for fruit borer damage in lychee fruit with aborted seeds is shown in Fig. 6(a). The AUC reached 0.982, demonstrating outstanding discrimination ability. Fig. 6 (b) presents the forward iPLS results showing the spectral intervals that were selected for classifying lychee fruit samples with aborted seed without and with fruit borer damage. The most relevant spectral regions for differentiating fruit borer damage in lychee fruit are primarily associated with variations in chemical composition, particularly sugars and moisture content. Absorption in the region around 4500 cm^{-1} can be attributed to the combination bands of the C—H bonds in sugars, while the 5500–6500 cm^{-1} interval is attributed to the first overtone of C—H stretching vibrations in CH_3 and CH_2 functional groups. Additionally, the region near 9000 cm^{-1} is associated with the second overtone of C—H stretching vibrations, also related to sugar content. The spectral range between 6900 and 7100 cm^{-1} corresponds to the first overtone of O—H stretching, representing water content. These spectral intervals are likely significant due to the impact of fruit borer activity, which can reduce both sugar and moisture levels, thereby altering the overall chemical composition of the fruit.

Table 6
Confusion matrices for the PLS-DA models based on variables selected using iPLS for detecting fruit borer damage in lychee fruit with aborted seeds: without fruit borer (AS w/o) and aborted seed with fruit borer (AS w) ($n_{\text{class AS w/o}} = 77$; $n_{\text{class AS}w} = 77$).

	A. Spectra stem end		B. Spectra at cheek side		C. Spectra at tip	
LVs	4		4		6	
	Actual		Actual		Actual	
	Class AS w/o	Class AS w	Class AS w/o	Class AS W	Class AS w/o	Class AS W
Predicted as Class AS w/o	69	6	71	12	63	10
Predicted as Class AS w	8	71	6	65	14	67
Predicted as unassigned	0	0	0	0	0	0

Table 7
Performance metrics for the test set of the PLS-DA models based on variables selected using iPLS from the NIR spectra acquired at different locations on the fruit for classification of lychee fruit with aborted seed as without fruit borer damage (AS w/o) and with fruit borer damage (AS w).

	A. Stem end			B. Cheek side			C. Tip		
	Sensitivity	Specificity	F1	Sensitivity	Specificity	F1	Sensitivity	Specificity	F1
Class AS w/o	0.90	0.92	0.90	0.73	0.98	0.83	0.82	0.87	0.84
Class AS w	0.92	0.90	0.91	0.98	0.73	0.87	0.87	0.82	0.85

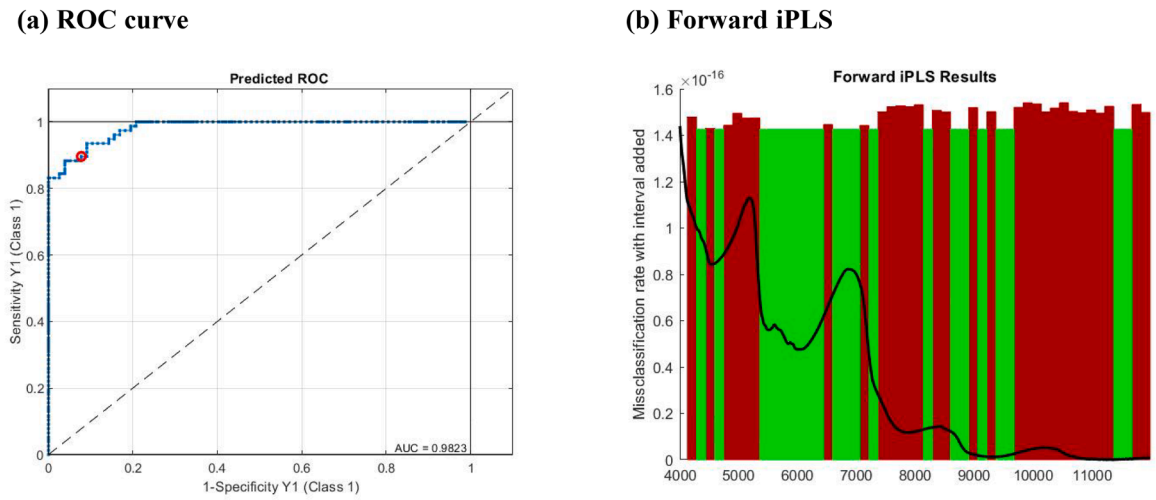


Fig. 6. (a) ROC curve and (b) Forward iPLS of classifying lychee fruit as aborted seed without and with worm using spectra at stem end.

3.3. Prediction of SSC and TA in lychee

The characteristics of the calibration and validation sample sets used for developing calibration models of Lychee fruit quality parameters are summarized in Table 8. The table presents the minimum, maximum, mean, and standard deviation (SD) of each measured parameter, providing an overview of the variability within the sample sets. In the calibration set ($n = 311$), SSC values and TA values respectively ranged from 15.2 to 21.6°Bx and from 0.08 to 0.50 %, while the ranges covered by the validation set were a bit smaller (15.7–20.8 °Bx and 0.15–0.42 %). These statistics indicate that the sample sets encompass a representative range of the quality parameters, which is essential for developing and evaluating calibration models.

The performance metrics for predicting the soluble solids content (SSC) and titratable acidity (TA) of lychee fruit based on the NIR spectra acquired at different locations on fruit without fruit borer damage with normal seed size and aborted seeded size are reported in Table 9. The best performance in cross-validation for SSC prediction ($R^2_p = 0.911$, RRMSEP = 2.09 %) was obtained by a PLSR model with 10 LVs using the spectra acquired at the cheek side (B). This might be attributed to the larger area of lychee flesh sampled at the cheek side compared to the stem end and tip positions. Nevertheless, all measurement positions gave $R^2_p > 0.85$, suggesting that SSC results are applicable for predicting lychee quality in laboratory testing (Porep et al., 2015). These results are consistent with those reported by Reddy & Yadav (2018). However, Similar trends were found for titratable acidity (TA) where the spectra acquired at the cheek side position also gave the highest R^2 value (0.901) with RRMSEP of 7.60 %, compared to those acquired at the stem end and apex positions ($R^2 = 0.890$ and 0.896, respectively). In Fig. 7(a) and (c), the SSC and TA values predicted based on NIR spectra acquired at the cheek side are plotted against the measured values. In line with the higher mentioned performance metrics, the samples are close to the bisectrix (1:1 line) in the scatter plot (Duarte et al., 2022).

The regression coefficient spectra of the PLSR models for predicting the SSC and TA values of lychee fruit from the NIR spectra are plotted in Fig. 7(b) and (d), respectively. The observed peaks at 5918, and 8100 cm^{-1} are associated with the soluble solids, primarily sugars, found in lychee fruit and can be attributed to a combination of C—H stretching and O—H stretching in carbohydrates (González-Caballero et al. 2010; Oliveira-Folador et al. 2018; Sringarm et al. 2024b). The region between 10,288 cm^{-1} corresponds to the third overtone of CH_2 stretching vibrations of the soluble solids. The regression coefficient spectra for predicting the TA value show peaks in the 5736–6000 cm^{-1} range, which are associated with the stretching vibration of the $C = O$, a functional group present in the acids (Oliveira-Folador et al., 2018). Additionally, an absorption band appears around 4500–4600 cm^{-1} which can be attributed to the combination band of O—H in organic acid groups (Ye et al., 2014).

Table 8
Characteristics of calibration and validation sample sets for the establishment of calibration models for the prediction of quality parameters of Lychee fruits from the acquired NIR spectra.

Parameters	Calibration set ($n = 311$)				Validation set ($n = 155$)			
	Mean	Min	Max	SD	Mean	Min	Max	SD
SSC (°Bx)	19.5	15.2	21.6	1.29	17.9	15.7	20.8	1.24
TA (%)	0.30	0.08	0.50	0.1	0.3	0.15	0.42	0.06

(A) stem end, (B) cheek side, (C) tip positions.

Table 9
Performance metrics for the prediction of the flavour parameters from the NIR spectra acquired at different locations on the lychee fruit.

Parameters	Measured positions	LVs	R^2_{cv}	RMSECV	R^2_p	RMSEP	RPD
SSC (°Bx)	A	10	0.902	0.36	0.899	0.41	3.02
	B	10	0.938	0.33	0.911	0.38	3.23
	C	9	0.917	0.38	0.902	0.40	3.10
TA (%)	A	10	0.898	0.03	0.890	0.02	2.87
	B	10	0.917	0.02	0.901	0.02	3.16
	C	10	0.900	0.03	0.896	0.02	3.01

(A) stem end, (B) cheek side, (C) tip positions.

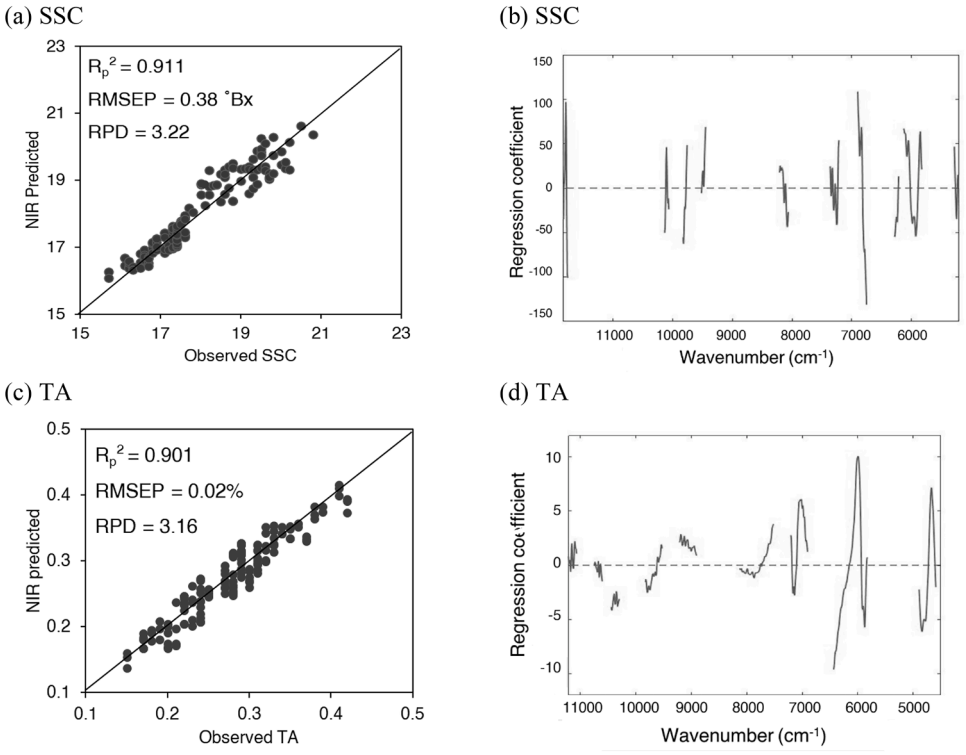


Fig. 7. Scatter plots of predicted versus measured values (left) for (a and b) soluble solids content (SSC), (c and d) Titratable acidity (TA), and corresponding regression coefficients (right).

It should be noted that the present study was conducted using a single lychee cultivar ('Pa Chid') harvested from one region and during two production seasons. The classification and regression models developed may be influenced by cultivar-specific characteristics such as peel thickness, flesh texture, and pigment content, as well as by environmental conditions including temperature, humidity, and growing practices. Therefore, further validation across multiple cultivars and production regions is recommended to verify if the models are robust against this variation or have to be updated to cope with this.

4. Conclusion

This study demonstrated the potential of Near Infrared (NIR) spectroscopy as a non-destructive and rapid method for grading lychee fruits based on their seed size and fruit borer damage, as well as for quantifying their soluble solids content (SSC) and titratable acidity (TA). The results indicate that NIR spectra acquired at the tip side of the fruit are most informative for distinguishing fruit with normal and aborted seeds with a balanced accuracy of 90.58 %, whereas spectra acquired from the stem end were most suitable for detecting fruit borer damage, with classification accuracies of 85.71 % and 90.91 % for normal and aborted seeds, respectively. The prediction models for SSC and TA using the spectra acquired at the cheek side demonstrated good accuracy, with R^2 values greater than 0.90, suggesting their suitability for quantitative analysis in both laboratory and industrial settings. Given the high classification accuracy and non-destructive nature of the technique, these findings suggest potential for integration into automated grading lines for lychee processing. However, it should be noted that different measurement locations were optimal for predicting the different quality traits, as the penetration of NIR radiation into lychee fruit is limited. This consideration should be addressed in the design of industrial grading systems. While this implies that three separate NIR measurements, at the tip, cheek, and stem end may be required per fruit to fully assess all quality parameters with the highest accuracy. This might be feasible for high-throughput applications if integrated into an automated system. The added complexity could be justified by the improved grading precision and comprehensive quality profiling enabled by this approach.

Ethical statement

Ethical approval was not required as this study did not include any experiments involving humans or animals.

Ethical statement

Ethical approval was not required as this study did not include any experiments involving humans or animals.

CRediT authorship contribution statement

Saowaluk Rungchang: Writing – original draft, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Wouter Saeys:** Writing – review & editing, Supervision, Methodology, Formal analysis, Conceptualization. **Chayanid Sringarm:** Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Sonthaya Numthuam:** Writing – review & editing, Methodology, Conceptualization. **Suphannika Intanon:** Writing – review & editing, Software, Formal analysis. **Sila Kittiwachana:** Writing – review & editing, Software, Formal analysis. **Sudarat Jiamyangyuen:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank B.N. Farm Co., Ltd., Thailand, for providing the lychee fruits used in this study. This work was supported by Chiang Mai University and the Program Management Unit for Competitiveness (PMUC) (Grant number C02F660010).

Data availability

Data will be made available on request.

References

Bázár, G., Romvári, R., Szabó, A., Somogyi, T., Éles, V., & Tsenkova, R. (2016). NIR detection of honey adulteration reveals differences in water spectral pattern. *Food Chemistry, 194*, 873–880. https://doi.org/10.1016/j.foodchem.2015.08.092

Bhushan, B., Pal, A., Narwal, R., Meena, V. S., Sharma, P. C., & Singh, J. (2015). Combinatorial approaches for controlling pericarp browning in Litchi (*Litchi chinensis*) fruit. *Journal of Food Science and Technology, 52*, 5418–5426. https://doi.org/10.1007/s13197-015-1712-8

Chu, Y., Lin, T., & Chang, J. (2015). Pollen effects on fruit set, seed weight, and shriveling of '73-S-20' Litchi-with special reference to artificial induction of parthenocarpy. *HortScience Horts, 50*(3), 369–373. https://doi.org/10.21273/HORTSCI.50.3.369. Retrieved Sep 2, (2024), from.

de Oliveira, G. A., de Castilhos, F., Renard, C. M. G. C., & Bureau, S. (2014). Comparison of NIR and MIR spectroscopic methods for determination of individual sugars, organic acids and carotenoids in passion fruit. *Food Research International (Ottawa, Ont.), 60*, 154–162. https://doi.org/10.1016/j.foodres.2013.10.051

Dong, X., Dong, Y., Liu, J., Wang, C., Bao, C., Wang, N., Zhao, X., & Chen, Z. (2024). Identification and quantitative detection of illegal additives in wheat flour based on near-infrared spectroscopy combined with chemometrics. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 323*, Article 124938. https://doi.org/10.1016/j.saa.2024.124938

Dong, X., Dong, Y., Liu, J., & Wu, S. (2025). Combined use of near infrared spectroscopy and chemometrics for the simultaneous detection of multiple illicit additions in wheat flour. *Applied Food Research, 5*(2), Article 101263. https://doi.org/10.1016/j.afres.2025.101263

Duarte, E. S., de, A., de Almeida, V. E., da Costa, G. B., de Araújo, M. C. U., Véras, G., Diniz, P. H. G. D., Fernandes, D. D., & de, S. (2022). Feasibility study on quantification and authentication of the cassava starch content in wheat flour for bread-making using NIR spectroscopy and digital images. *Food Chemistry, 368*. https://doi.org/10.1016/j.foodchem.2021.130843

González-Caballero, V., Sánchez, M. T., López, M. I., & Pérez-Marín, D. (2010). First steps towards the development of a non-destructive technique for the quality control of wine grapes during on-vine ripening and on arrival at the winery. *Journal of Food Engineering, 101*, 158–165. https://doi.org/10.1016/j.jfoodeng.2010.06.016

Goodarzi, M., Sharma, S., Ramon, H., & Saeys, W. (2015). Multivariate calibration of NIR spectroscopic sensors for continuous glucose monitoring. *TrAC - Trends in Analytical Chemistry, 67*, 147–158. https://doi.org/10.1016/j.trac.2014.12.005

Greensill, Colin V, & Walsh, Kerry B (2000). A remote acceptance probe and illumination configuration for spectral assessment of internal attributes of intact fruit. *Measurement Science and Technology, 11*, 1674. https://doi.org/10.1088/0957-0233/11/12/304

Kawano, S., Watanabe, H., & Iwamoto, M. (1992). Determination of sugar content in intact peaches by near infrared spectroscopy with Fiber optics in interactance mode. *Engei Gakkai zasshi. Engei Gakkai (Japan), 61*(2), 445–451. https://doi.org/10.2503/jjshs.61.445

Kumar, M., Kumar, V., Prasad, R., Varma, A. (2017). The lychee biotechnology, The Lychee Biotechnology. https://doi.org/10.1007/978-981-10-3644-6.

Lammertyn, J., Peirs, A., De Baerdemaeker, J., & Nicolaï, B. (2000). Light penetration properties of NIR radiation in fruit with respect to non-destructive quality assessment. *Postharvest Biology and Technology, 18*, 121–132. https://doi.org/10.1016/S0925-5214(99)00071-X

Li, S., Zhang, X., Shan, Y., Su, D., Ma, Q., Wen, R., & Li, J. (2017). Qualitative and quantitative detection of honey adulterated with high-fructose corn syrup and maltose syrup by using near-infrared spectroscopy. *Food Chemistry, 218*, 231–236. https://doi.org/10.1016/j.foodchem.2016.08.105

Lu, R., Van Beers, R., Saeys, W., Li, C., & Cen, H. (2020). Measurement of optical properties of fruits and vegetables: A review. *Postharvest Biology and Technology, 159*, Article 111003. https://doi.org/10.1016/j.postharvbio.2019.111003

Magwaza, L. S., & Opara, U. L. (2015). Analytical methods for determination of sugars and sweetness of horticultural products-A review. *Scientia Horticulturae, 184*, 179–192. https://doi.org/10.1016/j.scienta.2015.01.001

Manley, M. (2014). Near-infrared spectroscopy and hyperspectral imaging: Non-destructive analysis of biological materials. *Chemical Society Reviews, 43*, 8200–8214. https://doi.org/10.1039/c4cs00062e

Masithoh, R. E., Haff, R., & Kawano, S. (2016). Determination of soluble solids content and titratable acidity of intact fruit and juice of satsuma Mandarin using a hand-held near infrared instrument in transmittance mode. *Journal of Near Infrared Spectroscopy, 24*, 83–88. https://doi.org/10.1255/jnirs.1196

Matsuda, H., & Higuchi, H. (2019). Effects of the pollen parent on the fruit set and seed type of 'Bengal', 'Chakrapat', and 'Tai So' Lychee cultivars. *Tropical Agricultural Development, 63*(2), 79–82. https://doi.org/10.11248/jsta.63.79

Melado-Herreros, A., Nieto-Ortega, S., Olabarrieta, I., Gutiérrez, M., Villar, A., Zufía, J., Gorretta, N., & Roger, J. M. (2021). Postharvest ripeness assessment of 'Hass' avocado based on development of a new ripening index and vis-NIR spectroscopy. *Postharvest Biology and Technology*, 181. https://doi.org/10.1016/j.postharvbio.2021.111683

Nascimento, P. A. M., Carvalho, L. C.de, Júnior, L. C. C., Pereira, F. M. V., & Teixeira, G. H.de A (2016). Robust PLS models for soluble solids content and firmness determination in low chilling peach using near-infrared spectroscopy (NIR). *Postharvest Biology and Technology, 111*, 345–351. https://doi.org/10.1016/j.postharvbio.2015.08.006

Nicolai, B., Defraeye, T., De Ketelaere, B., Herremans, E., Hertog, M., Saeys, W., Toricelli, A., Vandendriessche, T., & Verboven, P. (2014). Nondestructive measurement of fruit and vegetable quality. *Annual Review of Food Science and Technology, 5*, 285–312. https://doi.org/10.1016/j.postharvbio.2020.111246

Nørgaard, L., Saudland, A., Wagner, J., Nielsen, J. P., Munck, L., & Engelsen, S. B. (2000). Interval partial least-squares regression (iPLS): A comparative chemometric study with an example from near-infrared spectroscopy. *Applied Spectroscopy, 54*(3), 413–419. https://doi.org/10.1366/0003702001949500

Oliveira-Folador, G., Bicudo, M., de, O., de Andrade, E. F., Renard, C. M. G. C., Bureau, S., & de Castilhos, F. (2018). Quality traits prediction of the passion fruit pulp using NIR and MIR spectroscopy. *Lwt, 95*, 172–178. https://doi.org/10.1016/j.lwt.2018.04.078

Ong, P., Chen, S., Tsai, C. Y., Wu, Y. J., & Shen, Y. T. (2022). A non-destructive methodology for determination of cantaloupe sugar content using machine vision and deep learning. *Journal of the Science of Food and Agriculture, 102*, 6586–6595. https://doi.org/10.1002/jsfa.12024

Porep, J. U., Kammerer, D. R., & Carle, R. (2015). On-line application of near infrared (NIR) spectroscopy in food production. *Trends in Food Science & Technology, 46*, 211–230. https://doi.org/10.1016/j.tifs.2015.10.002

Raz, A., Goldway, M., Sapir, G., & Stern, R. A (2022). Hong Long" Lychee (*Litchi chinensis* Sonn.) is the optimal pollinizer for the main lychee cultivars in israel. *Plants, 11*, 1996. https://doi.org/10.3390/plants11151996

Reddy, N., & Yadav, B. (2018). Nondestructive analysis of litchi fruit quality using FT-NIR spectroscopy. *International Journal of Biochemistry Research & Review, 22*, 1–10. https://doi.org/10.9734/ijbcrr/2018/41703

Saeys, W., Nguyen, D., Van Beers, R., & Nicolai, B. (2019). Multivariate calibration of spectroscopic sensors for postharvest quality evaluation: A review. *Postharvest Biology and Technology, 158*, Article 110981. https://doi.org/10.1016/j.postharvbio.2019.110981

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. https://doi.org/10.1016/j.postharvbio.2003.08.007

Sringarm, C., Numthuam, S., Singanusong, R., Jiamyangyuen, S., Kittiwatchana, S., Funsueb, S., & Rungchang, S. (2022). Quantitative determination of quality control parameters using near infrared spectroscopy and chemometrics in process monitoring of tapioca sweetener production. *Lwt*, 167. https://doi.org/10.1016/j.lwt.2022.113876. August.

Sringarm, C., Numthuam, S., Jiamyangyuen, S., Klangpetch, W., Wongsaipun, S., Kittiwachana, S., Saeys, W., & Rungchang, S. (2024a). Quantification of individual sugars in tapioca syrups with near-infrared spectroscopy. *Journal of Food Composition and Analysis, 125*, Article 105852. https://doi.org/10.1016/j.jfca.2023.105852

Sringarm, C., Numthuam, S., Jiamyangyuen, S., Kittiwachana, S., Kielar, F., Wongsaipun, S., & Rungchang, S. (2024b). *Quantitative and qualitative evaluation of maltodextrin products in the industry using near-infrared spectroscopy*. 1–12. https://doi.org/10.1111/ijfs.17476.

Sun, C., Aernouts, B., & Saeys, W. (2023). Characterisation and optical detection of puffy Satsuma mandarin. *Biosystems Engineering, 229*, 18–31. https://doi.org/10.1016/j.biosystemseng.2023.03.011

Theanjumpol, P., Wongzeewasakun, K., Muenmanee, N., Wongsaipun, S., Krongchai, C., Changrue, V., Boonyakiat, D., & Kittiwachana, S. (2019). Non-destructive identification and estimation of granulation in 'sai Num Pung' tangerine fruit using near infrared spectroscopy and chemometrics. *Postharvest Biology and Technology, 153*, 13–20. https://doi.org/10.1016/j.postharvbio.2019.03.009

Wang, Y., & Liu, J. (2024). Fast detection of ammonia nitrogen concentration in biogas fluid using near-infrared spectroscopy combined with wavelength selection by coronavirus herd immunity optimizer. *Microchemical Journal, 207*, Article 112203. https://doi.org/10.1016/j.microc.2024.112203

Wedding, B. B., Wright, C., Grauf, S., & White, R. D. (2024). Wavelength variation of the depth of penetration of near infrared radiation in "Hass" avocado fruit. *Technology in Horticulture, 4*, 1–4. https://doi.org/10.48130/tihort-0024-0005

Xie, L., Ye, X., Liu, D., & Ying, Y. (2009). Quantification of glucose, fructose and sucrose in bayberry juice by NIR and PLS. *Food Chemistry, 114*(3), 1135–1140. https://doi.org/10.1016/j.foodchem.2008.10.076

Xie, H., Wang, D., Qin, Yaqi, Ma, A., Fu, J., Qin, Yonghua, Hu, G., & Zhao, J. (2019). Genome-wide identification and expression analysis of SWEET gene family in *Litchi chinensis* reveal the involvement of LcSWEET2a/3b in early seed development. *BMC Plant Biology, 19*, 1–13. https://doi.org/10.1186/s12870-019-2120-4

Ye, M., Yue, T., Yuan, Y., & Li, Z. (2014). Application of FT-NIR spectroscopy to apple wine for rapid simultaneous determination of soluble solids content, pH, total acidity, and total ester content. *Food Bioprocess Technol, 7*, 3055–3062. https://doi.org/10.1007/s11947-014-1385-8

Zhang, Y., Huber, D. J., Hu, M., Jiang, G., Gao, Z., Xu, X., Jiang, Y., & Zhang, Z. (2018). Delay of postharvest browning in litchi fruit by melatonin via the enhancing of antioxidative processes and oxidation repair. *Journal of Agricultural and Food Chemistry, 66*, 7475–7484. https://doi.org/10.1021/acs.jafc.8b01922