



# Improving the non-destructive maturity classification model for durian fruit using near-infrared spectroscopy

Sirirak Ditcharoen<sup>a</sup>, Panmanas Sirisomboon<sup>b</sup>, Khwantri Saengprachatanarug<sup>a</sup>, Arthit Phuphaphud<sup>a</sup>, Ronnarit Rittiron<sup>c</sup>, Anupun Terdwongworakul<sup>d</sup>, Chayuttapong Malai<sup>a</sup>, Chirawan Saenphon<sup>a</sup>, Lalita Panduangnate<sup>a</sup>, Jetsada Posom<sup>a,\*</sup>

<sup>a</sup> Department of Agricultural Engineering, Faculty of Engineering, Khon Kaen University, Khon Kaen 40002, Thailand

<sup>b</sup> Department of Agricultural Engineering, School of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand

<sup>c</sup> Department of Food Engineering, Faculty of Engineering at Kamphaengsaen, Kasetsart University, Nakhon Pathom 73140, Thailand

<sup>d</sup> Department of Agricultural Engineering, Faculty of Engineering at Kamphaeng Saen, Kasetsart University, Kamphaeng Saen, Nakhon Pathom 73140, Thailand

## ARTICLE INFO

### Article history:

Received 10 October 2022

Received in revised form 13 February 2023

Accepted 19 February 2023

Available online 22 February 2023

### Keywords:

Classification

Durian fruit

Maturity stage

Near-infrared spectroscopy

Non-destructive method

## ABSTRACT

The maturity state of durian fruit is a key indicator of quality before trading. This research aims to improve the near-infrared (NIR) model for classifying the maturity stage of durian fruit using a completely non-destructive measurement. Both NIR spectrometers were investigated: the short wavelength NIR (SWNIR) ranging from 450 to 1000 nm and long wavelength NIR (LWNIR) ranging from 860 to 1750 nm. The samples collected for experimentation consisted of four stages: immaturity, prematurity, maturity, and ripe. Each fruit was scanned at the rind position on the main fertile lobe (header, middle, and tail) and stem. The classification models were developed using three supervised machine learning algorithms: linear discriminant analysis (LDA), support vector machine (SVM), and K-Nearest neighbours (KNN). The analysis results revealed that the use of durian rind spectra only obtained between 83.15% and 88.04% accuracy for the LWNIR spectrometer, while the SWNIR spectrometer provided 64.73 to 93.77% accuracy. The performance of model increases when developing with combination between rind and stem spectra. The LDA model developed using a combination of rind and stem spectra provided the greatest efficiency, exhibiting 97.28% and 100% accuracy for LWNIR and SWNIR spectrometers, respectively. The LDA model is therefore recommended for obtaining spectra from smoothing moving average (MA) + baseline of rind position and when used in combination with the MA + standard normal variance (SNV) of stem spectra. The NIR spectroscopy indicated high potential for non-destructive estimation of the durian maturity stage. This process could be used for quality control in the durian export industry to solve the problem of unripe durian being mixed with ripe fruit.

© 2023 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Durian is known as the king of fruit and considered to be an important agricultural product. Thailand is the largest producer and exporter of durian in the world (Ministry of Commerce, 2020). In particular, the Mon Thong variety is famous because it has a sweet, delicious taste. According to Thailand's export statistics for fresh durian in 2021, Thailand is the world's number one exporter of durian, with an export volume of around 900,000 tons, valued at about 3 billion dollars. In 2021, the value of Thailand's fresh durian exportation increased by approximately 66.39% compared to 2020 (Export statistics of fresh durian, 2022).

Presently, market demand for durian fruit is showing a continuously increasing trend, leading to high-value exports.

The quality of durian exports is controlled under the sorting criteria of the Agricultural Commodity Standards of the National Bureau of Agricultural Commodity and Food Standards, Ministry of Agriculture and Cooperatives to retain high quality and meet the specified product standards (Product Standard of Durian Agriculture, 2013). The quality of durian fruit for export is based on the analysis of its physical properties and external characteristics such as shape, amount of pulp, cleanliness without contamination, external defects, etc. (Somsri, 2013). For example, to be considered a mature fruit, durian must be harvested when its rind and stem colour change from bright green to brown or dark green. However, there is a high possibility that mistakes may be made if the fruit grows on the outside branches of the tree (under sunlight). Fruit exposed to extreme sunlight tends to be browner, rougher, and have

\* Corresponding author.

E-mail address: [jetspo@kku.ac.th](mailto:jetspo@kku.ac.th) (J. Posom).

dark brown stems, while the end of the thorn is drier and browner than canopied fruit of the same age (Regional information warehouse, 2023). Therefore, the traditional visual sorting method is subject to potential mistakes.

Most of the durian fruit exported is fresh. To deliver fresh fruit, it is necessary for the fruit to be transported under the appropriate conditions, such as by considering the transportation distance. If the transportation distance is short, mature durians could be chosen which can ripen within 1–3 days, while a longer transportation distance requires durians to be at a less mature stage to allow enough time for the fruit to be ripe and ready to eat when it reaches the consumer. Usually, durians are transported under refrigeration to retain the appropriate temperature (Paomnuaywit, 2019).

According to a recent report, from 2014 to 2015, the value of durian exports reduced from 69.3 to 6.5% because the quality of the fruit failed to meet the required standards (Jintakanon, 2020). If this situation were to continue, it could affect the international export business sector, causing the loss of its existing customer base. If the durians delivered do not meet the required standard, they will be beaten and sold at a lower price causing high-value losses.

To control quality and reduce the problem of young durian being mixed with mature fruit, random selection is employed to determine its chemical values, such as total soluble solid content (TSS) and dry matter (DM), allowing the inside quality of the durian to be thoroughly and accurately assessed. This method also has limitations in terms of cost because such analysis is time-consuming and requires highly-skilled operators. Therefore, it is not popularly used and cannot guarantee 100% quality.

Currently, near infrared (NIR) is used as an alternative method and has the potential to predict the quality of durian fruit. Talabnark and Terdwongworakul (2017) used FT-NIR spectroscopy ( $12,500\text{--}3,500\text{ cm}^{-1}$ ) to assess the DM content of durian pulp. The PLSR model developed from pulp spectra and whole fruit spectra with sliced rind spectra were compared for performance. The pulp spectra model provided an  $r_p$  and RMSEP of 0.78% and 3.02%, respectively, while the model developed from whole fruit with sliced rind spectra gave an  $r_p$  and RMSEP of 0.69% and 3.57%, respectively. Somton et al. (2015) applied near-infrared spectroscopy (1100–2500 nm) to classify three maturity classes (immature durian, early-mature durian, and mature durian) based on sliced rind and sliced stem absorbance. They concluded that a combination model using both sliced rind and stem spectra data provided the highest classification accuracy (94.4%). The researchers in the foregoing studies employed minimally destructive evaluation by removing the stems and slicing the peel and rind to provide a flat area before scanning. However, since fruit for export should not be scratched or damaged, this application is impractical. To address this research gap, an effective non-destructive test should be developed in the future.

Near-infrared spectroscopy was used to predict the sweetness of durian pulp both online (scanning with the sample moving) and offline (scanning with the sample remaining static) (Leepaitoon, 2018). The online model consisted of an  $R^2$  of 0.70 and RMSEP of 4.0%. Meanwhile, the offline model had an  $R^2$  of 0.81 and RMSEP of 3.27%. The near-infrared (NIR) hyperspectral image technique was used to scan the durian pulp, predict DM, and classify the ripeness of the durian pulp (Sharma et al., 2022). For DM prediction, the model developed with a genetic algorithm (GA) for variable selection provided the highest performance with an  $R_p^2$  of 0.878 and RMSEP of 1.22%. For ripeness classification, the LDA algorithm demonstrated 100% accuracy. Durian (*Durio zibethinus*) ripeness detection (unripe, ripe, and overripe) was performed using thermal image and multivariate analysis (PCA, KNN, LDA, SVM) (Mohd Ali et al., 2021). The classification accuracy was higher than 93% for each model. The SVM model exhibited the best performance, providing 97% accuracy. However, this research was conducted at room temperature.

According to the literature review, the near-infrared spectroscopy (NIRs) techniques demonstrated a high potential to determine the maturity stage and ripeness of durian fruit and pulp. The previous research resulted in destructive and minimally destructive fruit. Therefore, this present research attempted to perform evaluation under conditions which were as non-destructive as possible. Therefore, the aim of this study was to apply NIR spectroscopy techniques to classify the maturity of whole durian fruit. The scanning was performed under full non-destructive conditions and two main objectives: 1) to develop a classification model for measuring the maturity stage of whole durian fruit using SWNIR and LWNIR spectroscopy, and 2) to search for an optimal scanning method for use as a full non-destructive measurement of maturity class.

## 2. Materials and methods

### 2.1. Sample collection

Durian fruit (Mon Thong variety) was obtained from the Duangkhae durian orchard, Nam Nao District, Phetchabun Province (Latitude 16.7866, longitude 101.6478). A total of 160 fruit samples were collected from five different age periods, based on the days after blooming (DAB), including 80 to 90 (40 fruits), 90 to 100 (40 fruits), 100 to 110 (40 fruits), 110 to 120 (20 fruits), and 120 to 130 (20 fruits). The durian fruit collected was classified into three maturity stages based on DAB: unripe (80 to 100 DAB = 80 fruits), prematurity (101 to 120 DAB = 60 fruits), and mature (121 to 130 DAB = 20 fruits). The samples were taken to the Department of Agricultural Engineering Laboratory, Faculty of Engineering, KhonKaen University. The 30 durian fruits from prematurity stage (60 fruits) were separated randomly and were left to ripen for seven days. They were assigned as the ripe stage (Pimten., 2021; Posomboon, 2020). Durian is a climacteric fruit. The ripening is regulated by ethylene which fruit can ripe by itself after harvesting (Tongdee et al., 1990).

Therefore, the durian fruits of this experiment included four maturity stage (unripe: 80 fruits; prematurity: 30 fruits; mature: 20 fruits and ripe stage: 30 fruits). Table 1 shows the durian fruit categorised according to maturity and number of samples.

### 2.2. NIR scanning

The fruits sample were placed in room temperature at  $25 \pm 2^\circ\text{C}$  for one hour before the samples were taken for scanning to eliminate the temperature difference factor. The NIR scanning system consisted of a measurement chamber and two NIR spectrometers. The measurement chamber comprised a black box measuring  $60\text{ L} \times 60\text{ W} \times 60\text{ H cm}$ . Four halogen lamps of 25 W were installed at the top of each corner of the box at an angle  $45^\circ$  from the y-axis (see Fig. 1a). The samples were scanned using two spectrometers: long wavelength (LWNIR) spectrometer (AvaSpec-ULS2048-USB2-VA-50, AVANTES, Netherlands, wavelength range between 860 and 1750 nm), and short wavelength

**Table 1**  
Classification of ripening based on DAB and maturity percentage following to Thongklin (2018).

Stages	Maturity (%)	Day after blooming	Number of samples		
			Total (fruit)	Calibration set (fruit)	Validation set (fruit)
Unripe	$\leq 74$	80–100	80	61	19
Prematurity	75–80	101–120	30	21	9
Maturity	81–100	121–130	20	12	8
Ripe	$> 100$	leave durian to ripen	30	23	7

DAB: days after blooming.

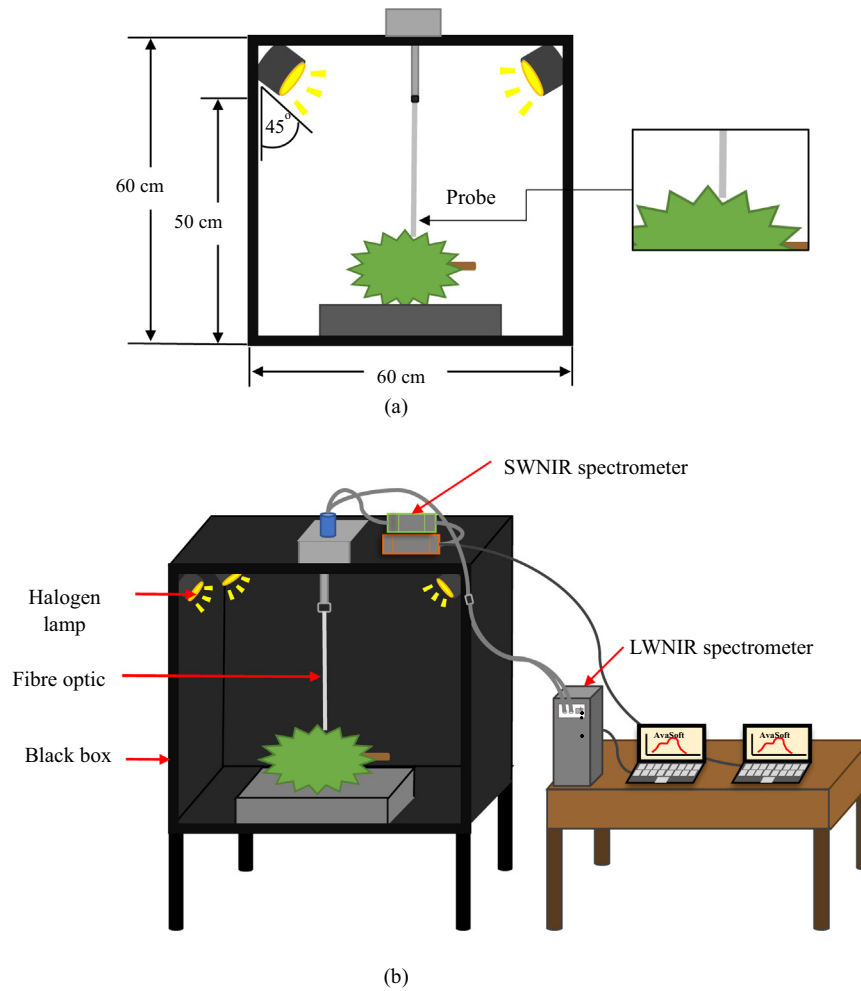


Fig. 1. Experimental setting, a) front view durian scanning, b) overview scanning and component of tools.

(SWNIR) spectrometer (AvaSpec-NIR256-1.7-EVO, AVANTES, Netherlands, wavelength range between 140 and 1100 nm).

Spectra acquisition consisted of three steps: 1) the light source was turned on and left for 10 min to ensure stable light intensity; 2) the dark and white references were collected; and 3) the samples were scanned using two spectrometers with the same integrating time of 200 ms. Each durian fruit sample was scanned under light windows. The reflectance was calculated as  $R = \frac{I_s - I_d}{I_w - I_d}$ , with  $I_s$ ,  $I_d$ , and  $I_w$  being the reflectance intensity of the sample, dark, and white reference, respectively. The absorbance ( $A$ ) was then calculated as  $A = \log(1/R)$ .

The samples were scanned at the stem (Somton et al., 2015) of one side in triplicate to obtain average spectra. The rind (at header, middle, and tail position) of the fertile lobe (which is a completing lobe throughout the length of the fruit and larger than other lobes) was then scanned in triplicate at each position and averaged (Buasub, 2007) (see Fig. 1b and Fig. 2). Therefore, four spectra were obtained for each fruit: header, middle, tail, and stem. The schematic of the developed classification model is shown in Fig. 2.

### 2.3. Dry matter content measurement

After scanning, the peel was removed from the fruit, and the pulp taken at the head, middle, and tail determined for DM using the dry oven method. The samples at each position weighed approximately 10 g according to the digital balance system (ATX224, Shimadzu, Japan) and placed in a hot air oven at 60 °C for 72 h (UF75, Memmert,

Germany) (Saechua et al., 2021). The dried samples were then weighed and the DM calculated following Eq. (1).

$$DM (\%) = \frac{W_2}{W_1} \times 100 \quad (1)$$

Where  $W_2$  is the final weight of the dried sample (g),  $W_1$  is the initial weight (g). Each position was evaluated in triplicate and averaged.

### 2.4. Model development

Fig. 2 displays the schematic for developing the model of maturity classification. To search for the best model, the total data were analyzed using supervised machine learning classification, linear discriminant analysis (LDA), support vector machine (SVM), and k-nearest neighbours (KNN). Spectral pre-treatment methods were used for both non-pre-processed (raw spectra) and pre-processed spectra from the 1st derivative (D1) (gap = 5 points, segment = 5 points), 2nd derivative (D2) (gap = 5 points, segment = 5 points), smoothing moving average (MA) + standard normal variance (Smooth + SNV), and MA + baseline offset. For the gap-segment derivative, the gap size indicated the spacing between points over which the derivative is computed, while the segment size indicated the range over which the points are averaged. Adding a gap and segment to a derivative can help to smooth and eliminate any noise created from the derivative itself.

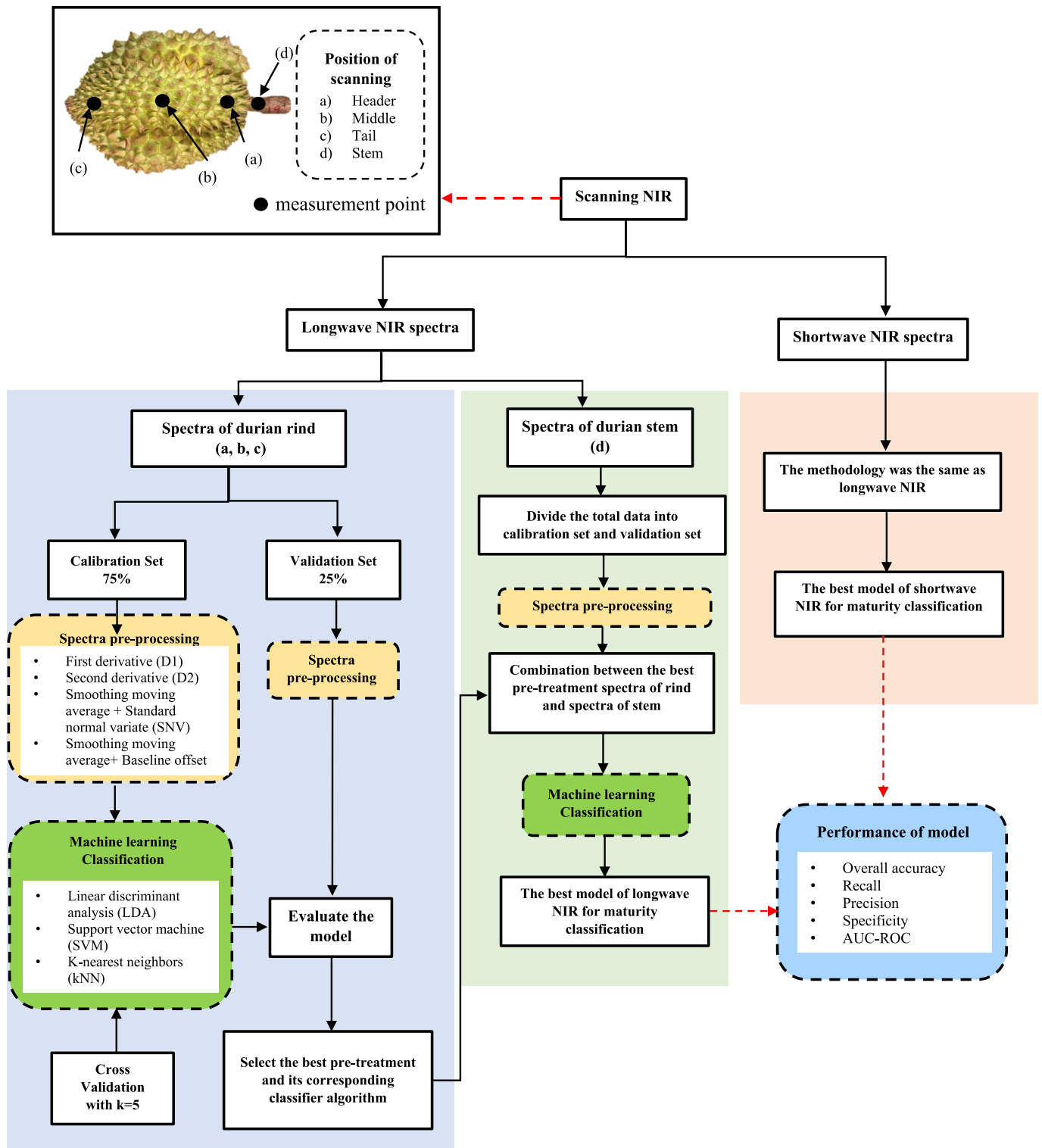


Fig. 2. Schematic for model development of maturity durian classification.

The training model was created using cross-validation (CV). K-Fold CV is where a given dataset is split into a K number of sections/folds with each fold used as a testing set at some point. Taking the scenario of 5-Fold CV ( $K = 5$ ). In the first iteration, the first fold is used to test the model and the remainder to train the model. In the second iteration, the second fold is used as the testing set, while the remainder serves as the training set. This process is repeated until each of the five folds have

been used as the testing set (Krishni, 2018). Thus, in this research, k-fold CV of  $k = 5$  used where the training set and test dataset are 75% and 25%, respectively. Therefore, the total sample was 160, calibration set was 117 fruits and validation set were 43 fruits.

In the first step, the models were developed using rind spectra (at positions: a (head), b (middle), and c (tail)), with each position assigned as an individual spectrum sample to choose the optimal

spectra pre-processing technique. Therefore, calibration set had 351 samples ( $117 \times 3$  spectra) and validation set had 129 samples ( $43 \times 3$  spectra). The rind spectra were examined with either raw or pre-processed spectra. The appropriate spectra pre-processing technique and its corresponding classification algorithm were then obtained.

In the second step, when the optimal spectra pre-processing method for rind (from first step) had been obtained, the pre-processed rind spectra were merged with raw stem spectra or stem spectra pre-treated with several techniques (D1, D2, MA + SNV, and MA + baseline offset). A rind spectral sample was merged with a stem sample to form a single sample. These combined spectra were used to generate the classification model again.

The performance of LDA, SVM, and KNN models was evaluated for overall accuracy (Eq. (2)), precision (Eq. (3)), recall (Eq. (4)), and specificity (Eq. (5)). The variables in each equation are true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

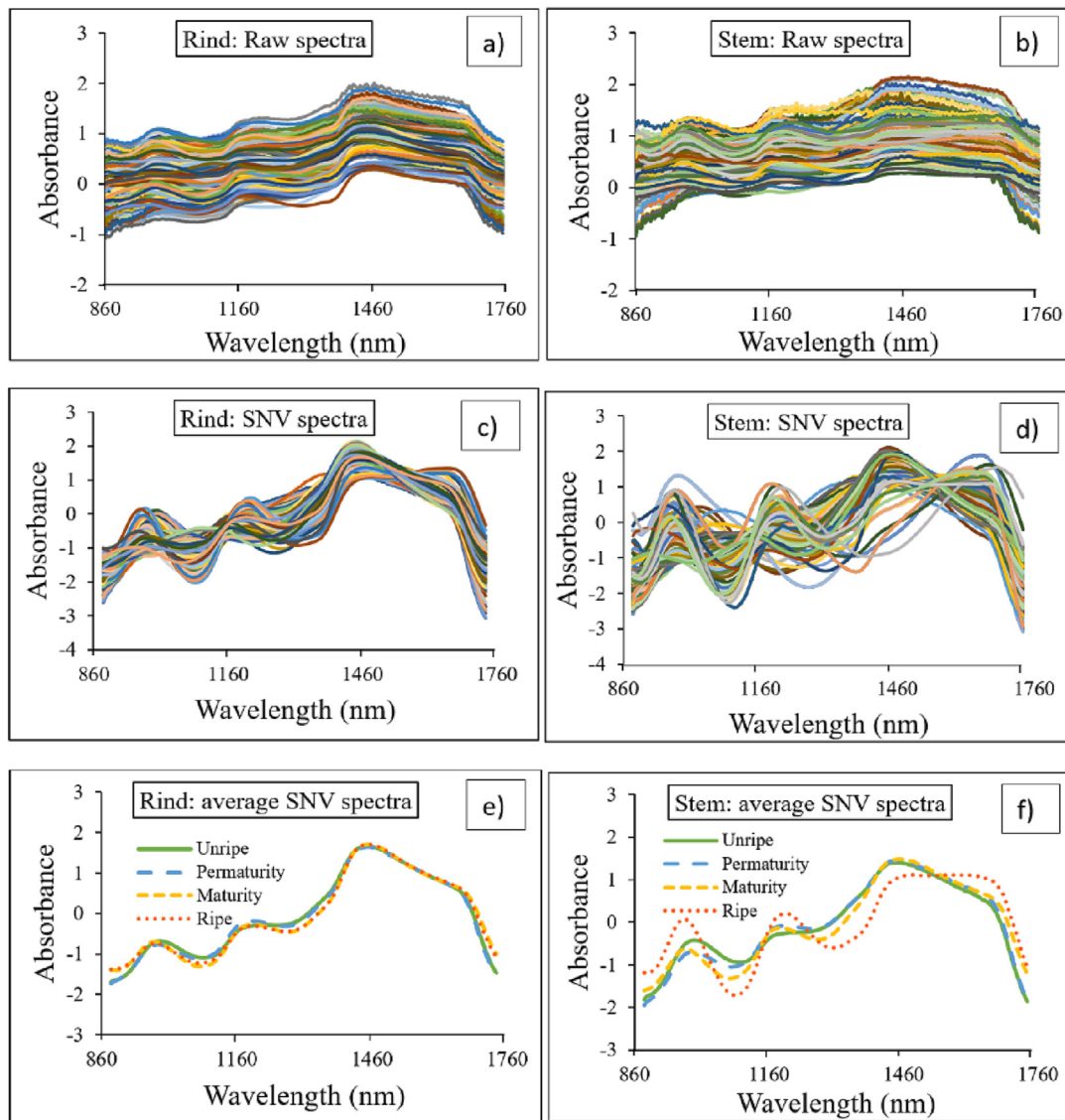
$$\text{Overall accuracy (\%)} = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (2)$$

$$\text{Precision (\%)} = \frac{TP}{TP + FP} \times 100 \quad (3)$$

$$\text{Recall (\%)} = \frac{TP}{TP + FN} \times 100 \quad (4)$$

$$\text{Specificity (\%)} = \frac{TN}{TN + FP} \times 100 \quad (5)$$

This work classifies into four classes i.e. unripe, prematurity, maturity, and ripe. For example, if considering for the class of ripe stage; TP, TN, FP, and FN can be described as: TP is the number of samples that are correctly classified for ripe; TN is the number of samples unripe, prematurity, and maturity; FN is the number of samples to unripe that are misclassified as ripe, prematurity, and maturity and; FP is the number of samples of unripe, prematurity, and maturity that are misclassified as ripe. Furthermore, the area under the receiver operating characteristic curve (AUC-ROC) estimated the potential of the classifier to distinguish between classes. The AUC-ROC ranges from 0 to 1, where a score close to 1 indicates that the classifier can successfully divide the positive class



**Fig. 3.** Spectra of LW NIR spectrometer for durian: (a) raw spectra of durian rind, (b) SNV spectra of durian rind, (c) average spectra of each maturity stage of durian rind, (d) raw spectra of durian stem, (e) SNV spectra of durian stem, and (f) average spectra of each maturity stage of durian stem.



from the negative. While remained three class was analyzed as same as above mention.

Both SWNIR and LWNIR spectra classification models were generated using the same methodology, and their performance subsequently compared. Finally, using different machine learning classification and pre-processing, the NIR wavelength spectrometer and spectra position model (only rind or combined spectra) provided the highest accuracy and was found to be the best model.

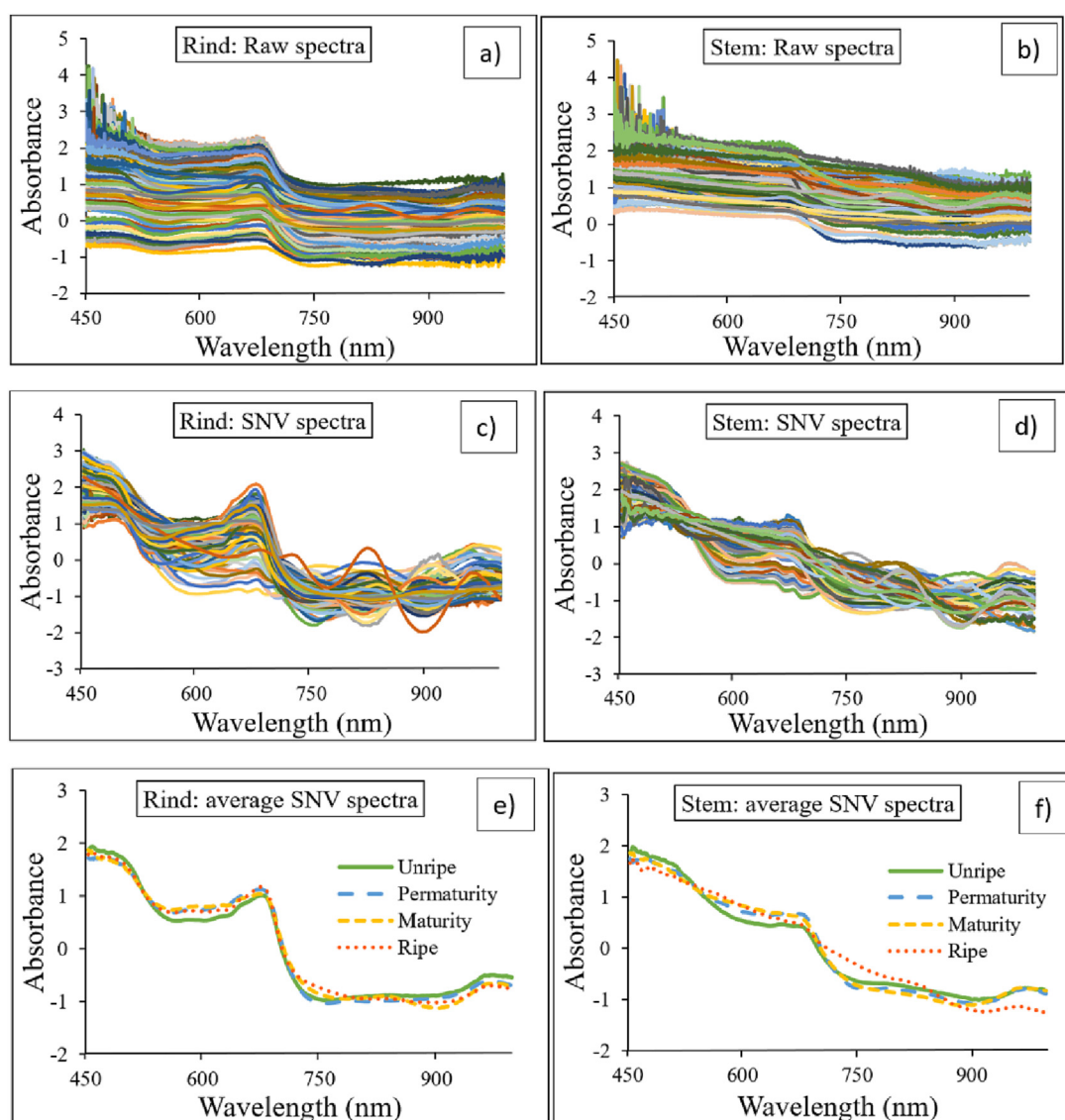
### 3. Results and discussion

#### 3.1. Spectral characteristics of durian

The raw spectra of durian rind and stem collected using LWNIR (860–1750 nm) and SWNIR (450–1000 nm) are shown in Figs. 3 and 4, respectively. Figs. 3a, b, and c; 4a, b, and c show the raw, SNV, and average raw spectra of the rind, respectively. Figs. 3d, e, a, and f; 4d, e, and f show the raw, SNV, and average raw spectra of the stem, respectively.

Figs. 3a, b, d, and e; and 4a, b, d, and e show the raw spectra and SNV spectra of the rind and stem, respectively. Baseline offset usually occurred when scanning the durian fruit samples. The linear baseline

shift affecting the absorption value was tilted up systematically with increasing wavelengths (Haruthaithanasan, 2011). The average raw spectrum of the rind and stem at different maturity stages, i.e., unripe, premature, mature, and ripe, are presented in Figs. 3c and f and 4c and f for LWNIR and SWNIR, respectively. The average raw spectra of both rind and stem were clearly divided into two groups. Group 1: unripe and premature stage and Group 2: mature and ripe stage, showing similar results. While the average raw spectra between the two groups was obviously different. For example, Fig. 3c and f show that the absorption of the two groups was clearly different at a wavelength of 1250 nm vibrational band of sugar (Pu et al., 2016). For SWNIR, the average raw rind spectra, at a wavelength of 680 nm vibrational band of anthocyanins and chlorophyll (Benelli et al., 2021), approximately, clearly had different absorption values at various stages. Whereas Fig. 4f shows that the SWNIR of raw stem spectra clearly had a different absorption curve. As previously mentioned, the different absorption values appeared clearly at various maturity stages resulting in the classification model giving good accuracy. The report by Sharma et al. (2022) shows the average spectra SNV pre-treated. At wavelengths of 900 and 1200 nm the spectra can be clearly split according to the different ripeness stages of durian. Likewise, Saputro and Handayan (2017)



**Fig. 4.** Spectra of SWNIR spectrometer for durian: (a) raw spectra of durian rind, (b) SNV spectra of durian rind, (c) average spectra of each maturity stage of durian rind, (d) raw spectra of durian stem, (e) SNV spectra of durian stem, and (f) average spectra of each maturity stage of durian stem.

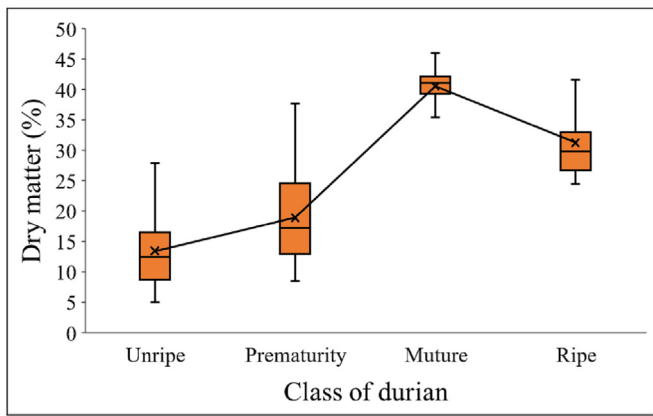


Fig. 5. Dry matter of durian pulp according to different maturity stage.

demonstrated that the reflectance profile of a banana was separated in the various maturity stages at wavelengths of around 480, 680, and 960 nm.

### 3.2. Dry matter content

Fig. 5 demonstrates the average DM of durian pulp at different maturity stages. The results show that the DM of durian pulp increased with a rise in DAB. However, durian at the ripe stage exhibited a lower DM than at the maturity stage because it was left to ripen naturally from the prematurity. These results are consistent with the findings of Chuenatsadongkot et al. (2018) in that the correlation of the DM percentage to the DAB shows a clear trend, with the DM percentage tending to increase steadily during the period, rising from 100 to 135 DAB. However, at the ripe stage, the mean percentage DM declined less than at the maturity stage because when the durian is very ripe,

the pulp smears, causing the humidity to be higher than usual. This corresponds to Timkhum and Terdwongworakul (2013), where the percentage DM values are lower because of the high moisture caused by the heavy rainfall. This demonstrates that humidity affects the percentage DM.

### 3.3. Classification models

Table 2 shows the accuracy of validation set of each classification model developed using the full wavelength from LWNIR. The report on the first step indicates that the maturity classification models were developed using durian rind spectra coupled with different algorithms (i.e., LDA, SVM, and KNN) and either raw spectra or pre-treated spectra (i.e., D1, D2, MA + SNV and MA + baseline offset) to find the optimal pre-treatment and its corresponding algorithm. The model efficiency was determined based on precision, recall, specificity, and accuracy. The results show that the optimal pre-treatment of the D2 was suitable for LDA and SVM algorithms, and the MA + SNV was appropriate for KNN, provided precision, recall, specificity, and overall accuracy for validation set of 71.45%, 71.62%, 91.61%, and 87.50%; 73.91%, 62.45%, 90.21%, and 88.04%; and 73.91%, 62.45%, 90.21%, and 88.04%, respectively. Finally, the optimal spectra pre-treatment of the D2 was appropriate for LDA and SVM, while MA + SNV was optimal for KNN.

According to the report on the second step, either the raw or pre-treated stem spectra were matched with the rind spectra (pre-treated using the D2 for LDA and SVM, and MA + SNV for KNN, obtained from the first step). From the hypothesis, it is clear that the combination of rind spectra and stem spectra could improve the classification accuracy. Table 2 indicates that the combined spectra (rind and stem) can increase the model performance of all algorithms. The LDA model developed with the rind (D2) + stem (MA + SNV) provided the best classification efficiency, precision, recall, specificity, and overall accuracy of 96.40%, 91.88%, 97.55%, and 97.28%, respectively. The confusion matrix of validation set was displayed in Table 3.

Table 2

The accuracy of the different classification models for validation set developed using the full wavelength of LWNIR spectrometer.

Model	Spectra	Rind	Stem	N	Precision	Recall	Specificity	Accuracy 5-fold CV	Overall accuracy	AUC-ROC
LDA	Durian rind	Raw	–	129	59.83	62.03	88.16	77.04 ± 0.64	83.15	0.82
		1 <sup>st</sup> derivative	–	129	69.27	70.59	91.12	65.96 ± 2.00	86.96	0.77
		2 <sup>nd</sup> derivative	–	129	71.45	71.62	91.61	66.40 ± 1.34	87.50	0.75
		MA + SNV	–	129	66.34	65.73	88.68	76.00 ± 1.76	84.78	0.78
		MA + Baseline offset	–	129	61.93	64.06	88.01	73.18 ± 2.86	83.15	0.79
	Durian rind combination with stem	2 <sup>nd</sup> derivative	Raw	129	91.85	85.32	90.00	84.80 ± 1.08	97.83	0.88
			1 <sup>st</sup> derivative	129	94.04	92.60	93.65	84.60 ± 1.99	93.48	0.90
			2 <sup>nd</sup> derivative	129	95.74	93.79	95.32	83.14 ± 1.22	94.57	0.88
			MA + SNV	129	96.40	91.88	97.55	82.54 ± 1.19	97.28	0.86
			MA + Baseline offset	129	83.60	85.27	95.45	79.65 ± 2.97	93.48	0.85
			–	129	67.45	64.28	90.92	66.02 ± 2.39	87.50	0.82
			–	129	73.17	63.54	90.43	71.82 ± 1.69	88.04	0.85
SVM	Durian rind	Raw	–	129	67.45	64.28	90.92	66.02 ± 2.39	87.50	0.82
		1 <sup>st</sup> derivative	–	129	73.17	63.54	90.43	71.82 ± 1.69	88.04	0.85
		2 <sup>nd</sup> derivative	–	129	73.91	62.45	90.21	75.76 ± 1.23	88.04	0.87
		MA + SNV	–	129	68.06	58.19	87.97	71.82 ± 1.69	85.33	0.85
		MA + Baseline	–	129	60.19	55.45	88.34	75.76 ± 1.23	85.33	0.87
	Durian rind combination with stem	2 <sup>nd</sup> derivative	Raw	129	91.57	81.23	96.17	80.50 ± 1.00	95.11	0.95
			1 <sup>st</sup> derivative	129	90.91	79.57	95.63	84.32 ± 0.98	94.57	0.94
			2 <sup>nd</sup> derivative	129	87.24	76.15	94.61	86.00 ± 1.33	92.93	0.95
			MA + SNV	129	75.88	66.00	91.65	81.93 ± 2.72	89.67	0.92
			MA + Baseline offset	129	85.45	82.44	95.33	83.35 ± 0.96	93.48	0.94
			–	129	61.64	57.67	87.99	53.70 ± 1.25	85.33	0.71
			–	129	55.38	56.16	89.46	66.14 ± 2.78	84.78	0.76
KNN	Durian rind	Raw	–	129	61.64	57.67	87.99	53.70 ± 1.25	85.33	0.71
		1 <sup>st</sup> derivative	–	129	55.38	56.16	89.46	66.14 ± 2.78	84.78	0.76
		2 <sup>nd</sup> derivative	–	129	73.91	62.45	90.21	69.82 ± 2.81	88.04	0.75
		MA + SNV	–	129	58.67	57.80	88.17	63.42 ± 1.47	84.24	0.78
		MA + Baseline	–	129	54.70	55.42	88.10	63.18 ± 1.06	83.15	0.71
	Durian rind combination with stem	2 <sup>nd</sup> derivative	Raw	129	74.77	65.36	90.04	75.18 ± 1.45	87.50	0.82
			1 <sup>st</sup> derivative	129	83.16	78.87	94.11	81.70 ± 0.82	92.39	0.85
			2 <sup>nd</sup> derivative	129	47.19	48.99	85.80	68.88 ± 1.95	79.35	0.75
			MA + SNV	129	75.63	73.16	92.82	79.06 ± 1.64	90.22	0.83
			MA + Baseline offset	129	75.88	66.00	91.65	78.75 ± 1.04	89.67	0.83

<sup>N</sup>Number of sample for validation set; <sup>MA</sup>Smoothing moving average; <sup>SNV</sup>Standard normal variate.

**Table 3**

Confusion matrix of the LDA model developed with the rind (2nd derivative) + stem (MA + SNV) for LWNIR spectrometer.

		Prediction			
		Unripe	Prematurity	Maturity	Ripe
Actual	Unripe	57	0	0	0
	Prematurity	4	23	0	0
	Maturity	0	0	24	0
	Ripe	2	0	1	18

Table 4 shows the accuracy of validation set of each classification model developed using the full wavelength from SWNIR. The analysis methodology was the same as for LWNIR. The maturity classification models were developed using durian rind spectra. Each algorithm coupled with either raw or pre-processed spectra was used for model development to find the optimal pre-treatment and its corresponding algorithm. The model efficiency was also determined based on precision, recall, specificity, and overall accuracy. The results show that the optimal pre-treatment of the MA + baseline offset was suitable for LDA, the D1 for SVM, and the D2 for KNN, provided precision, recall, specificity, and overall accuracy of 84.70%, 83.48%, 85.30%, 84.39%; 95.99%, 91.99%, 95.56%, 93.77%; and 87.90%, 85.62%, 87.50%, 87.01%, respectively.

As seen in Table 4, it can be observed that the combined spectra (rind and stem) can increase the model performance for all algorithms. The LDA model developed with the rind (MA + baseline offset) + stem (MA + SNV) provided the best classification efficiency, obtaining precision, recall, specificity, and overall accuracy of 100%. Table 5 demonstrate the confusion matrix of validation set for SWNIR spectrometer which model ability was higher slightly than that of LWNIR spectrometer.

The NIR-HSI (NIR hyperspectral image) system (900–1600 nm) was used to create the ripeness classification model (unripe, ripe, and

**Table 5**

Confusion matrix of the LDA model developed with the rind (MA + Baseline offset) + stem (MA + SNV) for SWNIR spectrometer.

		Prediction			
		Unripe	Prematurity	Maturity	Ripe
Actual	Unripe	57	0	0	0
	Prematurity	0	27	0	0
	Maturity	0	0	24	0
	Ripe	0	0	0	21

override) following DAB. The model was developed with pulp durian spectra. The results display that the LDA model could be used to classify the durian maturity level, providing high accuracy of 100% (Sharma et al., 2022). Talabnark and Terdwongworakul (2017) reported that FT-NIR spectroscopy (800–2500 nm) can be used to classify the fruit maturity levels (at 101, 108, 115, and 122 days) using sliced rind spectra with a high accuracy level of 87.48%, and the model developed by PLS-DA. Whereas the model established from pulp durian spectra provided higher accuracy than the sliced rind spectra (92.5%). Somton et al. (2015) found that the classification model from NIR (1000–2500 nm) developed using a combination of sliced rind and sliced stem spectra can achieve higher accuracy than models created from only sliced rind spectra (94.4% accuracy). Timkhum and Terdwongworakul (2012) applied visible NIR spectroscopy (350–750 nm) to classify five maturity stages (ranging from 106 to 134 DAB) using sliced rind durian spectra. The model achieved the best classification accuracy (83.30%), while the model developed with sliced stem spectra provided higher accuracy than the sliced rind spectra (94.7%) (Saputro and Handayani, 2017). These results were supported by Praneenarat, who found that the durian stem contains liquid sugar compounds that coincide with sweetness. The chemical analysis revealed that when the durians began to ripen, the sucrose increased significantly while the glucose and fructose decreased (Praneenarat, 2021). Therefore, the higher accuracy

**Table 4**

The accuracy of the different classification models for validation set developed using the full wavelength of SWNIR spectrometer.

Model	Spectra	Rind	Stem	N	Precision	Recall	Specificity	Accuracy 5-fold CV	Overall accuracy	AUC-ROC
LDA	Durian rind	Raw	–	129	76.37	76.79	78.92	86.86 ± 1.43	77.86	0.88
		1 <sup>st</sup> derivative	–	129	81.38	80.73	82.55	86.44 ± 1.62	81.64	0.87
		2 <sup>nd</sup> derivative	–	129	74.62	74.01	77.29	86.32 ± 1.81	75.65	0.87
		MA + SNV	–	129	79.05	80.70	78.45	90.10 ± 1.24	79.57	0.87
		MA + Baseline offset	–	129	84.70	83.48	85.30	90.18 ± 1.62	84.39	0.88
	Durian rind combination with stem	MA + Baseline offset	Raw	129	97.80	93.18	97.70	94.48 ± 1.26	95.44	0.95
			1 <sup>st</sup> derivative	129	100.00	100.00	100.00	95.08 ± 1.06	100.00	0.95
			2 <sup>nd</sup> derivative	129	99.46	97.73	99.44	94.06 ± 1.65	98.59	0.95
			MA + SNV	129	100.00	100.00	100.00	96.36 ± 0.99	100.00	0.96
			MA + Baseline offset	129	99.46	97.73	99.44	93.70 ± 0.73	98.59	0.94
SVM	Durian rind	Raw	–	129	80.83	72.66	82.97	69.02 ± 2.07	77.81	0.84
		1 <sup>st</sup> derivative	–	129	95.99	91.99	95.56	89.90 ± 0.65	93.77	0.93
		2 <sup>nd</sup> derivative	–	129	93.97	92.63	94.31	89.02 ± 1.52	93.47	0.94
		MA + SNV	–	129	79.73	77.86	80.22	81.78 ± 0.96	79.04	0.94
		MA + Baseline offset	–	129	74.66	72.89	76.04	72.64 ± 1.59	74.47	0.84
	Durian rind combination with stem	1 <sup>st</sup> derivative	Raw	129	92.69	89.18	92.62	87.60 ± 1.63	90.90	0.97
			1 <sup>st</sup> derivative	129	96.52	95.35	96.11	94.16 ± 0.70	95.73	0.99
			2 <sup>nd</sup> derivative	129	98.44	98.81	98.33	94.76 ± 0.92	98.57	0.99
			MA + SNV	129	95.63	95.95	95.48	92.80 ± 1.05	95.71	0.99
			MA + Baseline offset	129	93.06	93.57	92.14	88.67 ± 0.96	92.86	0.97
KNN	Durian rind	Raw	–	129	70.51	53.72	75.74	60.46 ± 0.52	64.73	0.72
		1 <sup>st</sup> derivative	–	129	85.87	84.61	85.40	82.74 ± 2.12	85.01	0.87
		2 <sup>nd</sup> derivative	–	129	87.90	86.52	87.50	86.52 ± 1.24	87.01	0.88
		MA + SNV	–	129	80.43	69.88	77.94	72.76 ± 1.22	73.91	0.87
		MA + Baseline offset	–	129	66.45	67.66	64.83	65.94 ± 1.18	66.24	0.72
	Durian rind combination with stem	2 <sup>nd</sup> derivative	Raw	129	98.44	98.81	98.33	86.96 ± 1.96	98.57	0.89
			1 <sup>st</sup> derivative	129	97.92	98.33	97.73	91.80 ± 0.91	98.03	0.94
			2 <sup>nd</sup> derivative	129	97.92	98.33	97.73	95.70 ± 0.88	98.03	0.95
			MA + SNV	129	98.44	98.81	98.33	92.80 ± 1.05	98.57	0.95
			MA + Baseline offset	129	97.92	98.33	97.73	88.67 ± 0.96	98.03	0.92

<sup>N</sup>Number of sample for validation set; <sup>MA</sup>Smoothing moving average; <sup>SNV</sup>Standard normal variate.



provided by the durian stem spectra model accords with the results of the present study (see Tables 2 and 4), while the spectra of the stem combined with rind gave high accuracy. Some previous studies have resulted in destructive measurement, and others minimally destructive. The aim of the present research was to provide a completely non-destructive measurement. The NIR spectroscopy, combined with machine learning has the potential to be used as a non-destructive measurement.

The application of NIR spectroscopy could reduce operating time, and different positions of rind combined with stem spectra producing a similar performance. Any rind position combined with stem spectra can be used for prediction. The findings reveal that different rind spectra can give similar results. Accurate prediction was achieved by the LDA classification model developed from MA + baseline rind spectra (different positions, i.e., head, middle, and tail) and MA + SNV stem spectra for SWNIR spectrometer.

#### 4. Conclusion

This research demonstrates that the NIR spectroscopy technique has the potential to become a full non-destructive measurement. The model was generated using several machine learning techniques for classification of the maturity (four stages, i.e., immature, premature, mature, and ripe) based on durian DAB. The strength of this research is that the samples are not destroyed during the classification process. The LWNIR and SWNIR coupled with LDA, SVM, and KNN algorithm gave slightly different performance results. The developed LWNIR model using rind spectra provided overall accuracy of 87.50%, 88.04%, and 88.04%, respectively, while the developed SWNIR model using rind spectra gave overall accuracy of 84.39%, 93.77%, and 87.01%, respectively. The models improved when combined spectra (rind spectra + stem spectra) were used. The LDA model was the best classifier, providing 100% classification accuracy for SWNIR when the rind spectra were pre-treated with the MA + baseline offset and the stem spectra pre-treated using MA + SNV. The rind spectra at any position combined with the stem spectra is recommended for practical use. The NIR spectrometer could be applied to classify durian maturity. Any spectrometer instruments could be used as it offers the same high efficiency, but SWNIR is recommended because it is less expensive than LWNIR.

#### CRediT authorship contribution statement

**Sirirak Ditcharoen:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Panmanas Sirisomboon:** Validation, Funding acquisition. **Khwantri Saengprachatanarug:** Validation, Conceptualization. **Arthit Phuphaphud:** Validation, Writing – review & editing. **Ronnarit Rittiron:** Validation. **Anupun Terdwongworakul:** Validation. **Chayuttapong Malai:** Validation, Data curation. **Chirawan Saenphon:** Validation, Data curation, Methodology. **Lalita Panduangnate:** Validation, Data curation. **Jetsada Posom:** Conceptualization, Formal analysis, Data curation, Writing – review & editing, Supervision, Project administration, Funding acquisition.

#### Declaration of Competing Interest

None.

#### Acknowledgements

This work also supported by Research and Graduate Studies, Khon Kaen University, Thailand; Research Fund for Supporting Lecturer to Admit High Potential Student to Study and Research on His Expert Program Year 2021 from Graduate School, Khon Kaen University, Thailand; the Agricultural Research Development Agency (Public Organisation) [grant number CRP6405031580]. The authors would like to

acknowledge the Department of Agricultural Engineering, Faculty of Engineering, Khon Kaen University, for providing the equipment and location for the experiments and King Mongkut's Institute of Technology Ladkrabang for its financial support.

#### References

- Benelli, A., Cevoli, C., Fabbri, A., Ragni, L., 2021. Ripeness evaluation of kiwifruit by hyperspectral imaging. *Biosyst. Eng.* <https://doi.org/10.1016/j.biosystemseng.2021.08.009>.
- Buasub, W., 2007. Agricultural extension academic manual (Durian). Bureau Agric. Commodit. Promot. Manag. accessed 1 August 2022. [http://www.agriman.doe.go.th/home/t.n1/3fruit\\_Requirement/01\\_Durian.pdf](http://www.agriman.doe.go.th/home/t.n1/3fruit_Requirement/01_Durian.pdf).
- Chuenatsadongkot, T., Treemnuak, T., Treemnuak, K., 2018. Comparison of the ability to evaluation of 'Monthong' durian maturity using color value from spectrometer and image processing. *Int. J. Mech. Prod. Eng. Issue 6*. <http://iraj.in>.
- Export statistics of fresh durian, 2022. Office of Agricultural Economics. [http://impexp.oae.go.th/service/export.php?S\\_YEAR=2563&E\\_YEAR=2565&PRODUCT\\_GROUP=5252&PRODUCT\\_ID=4977&wf\\_search=&WF\\_SEARCH=Y](http://impexp.oae.go.th/service/export.php?S_YEAR=2563&E_YEAR=2565&PRODUCT_GROUP=5252&PRODUCT_ID=4977&wf_search=&WF_SEARCH=Y).
- Haruthaithanasan, V., 2011. NIR-Infraede Technology and Applications in Industries. <https://ebook.lib.ku.ac.th/ebook27/ebook/2015RG0071/>.
- Jintakanon, K., 2020. Solving the problem of maturity durian. Technol. Chaoban, accessed 1 August 2022. [https://www.technologychaoban.com/agricultural-technology/article\\_183405](https://www.technologychaoban.com/agricultural-technology/article_183405).
- Krishni, 2018, December 17. K-Fold Cross Validation. DataDrivenInvestor. <https://medium.datadriveninvestor.com/k-fold-cross-validation-6b8518070833>.
- Leepaitoon, K., 2018. On-line and off-line nir spectroscopy for measuring total soluble solids of durian (Durio zibethinus CV monthong). King Mongkut's Institute of Technology Ladkrabang. <http://ebook.lib.kmit.ac.th/assets/library/reader/reading/index.php>.
- Ministry of Commerce, 2020. Durian is the king of Thai fruits. Trade Policy and Strategy Office. [http://www.tpsa.moc.go.th/sites/default/files/thueriyn\\_240863.pdf](http://www.tpsa.moc.go.th/sites/default/files/thueriyn_240863.pdf).
- Mohd Ali, M., Hashim, N., Shahamshah, M.I., 2021. Durian (Durio zibethinus) ripeness detection using thermal imaging with multivariate analysis. *Postharvest Biol. Technol.* 176. <https://doi.org/10.1016/j.postharvbio.2021.111517>.
- Paomnuaywit, P., 2019. How to export durian to be fresh until the destination. Logist. Manager, accessed 1 August 2022. <https://logistics-manager.com/th/how-to-deliver-durian/>.
- Pimten., 2021, February 15. How to rapid ripen durian. Shopper's Cafe; Elsevier B.V. <https://www.shopat24.com/blog/trending/how-to-make-durian-ripen-quickly-tips-that-durian-necks-should-not-miss>.
- Posomboon, M., 2020, February 15. How to cutting durian for 5 days have not come off the bouquet. Kasetgo; Elsevier B.V. <https://kasetgo.com/t/topic/8839>.
- Praneenarat, T., 2021. Durian ripeness indicator from the sugar in its stem. Chulalongkorn University. <https://www.chula.ac.th/en/highlight/50923/>.
- Product Standard of Durian Agriculture, 2013. National Bureau of Agricultural Commodity and Food Standards. [http://e-book.acfs.go.th/Book\\_view/256](http://e-book.acfs.go.th/Book_view/256).
- Pu, H., Liu, D., Wang, L., Sun, D.W., 2016. Soluble solids content and pH prediction and maturity discrimination of lychee fruits using visible and near infrared hyperspectral imaging. *Food Anal. Methods* 9 (1), 235–244. <https://doi.org/10.1007/s12161-015-0186-7>.
- Regional information warehouse, 2023. Botanical characteristics. Agricultural Research Development Agency Retrieved 2 August 2022, from <https://www.arda.or.th/kasetinfo/south/durian/controller/index.php>.
- Saputro, A., Handayani, W. (2017). Wavelength Selection in Hyperspectral Imaging for Prediction Banana Fruit Quality. 2017 International Conference on Electrical Engineering and Informatics (ICELTICS 2017) October 18–20, 2017 - Banda Aceh, Indonesia. <https://doi.org/10.1109/ICELTICS.2017.8253259>.
- Saechua, W., Sharma, S., Nakawajana, N., Leepaitoon, K., Chunsri, R., Posom, J., Roeksukrungrueang, C., Siritechavong, T., Phansomophon, T., Sirisomboon, P., Lapcharoensuk, R., Pornchaloempong, P., 2021. Integrating Vis-SWNIR spectrometer in a conveyor system for in-line measurement of dry matter content and soluble solids content of durian pulp. *Postharvest Biol. Technol.* 181. <https://doi.org/10.1016/j.postharvbio.2021.111640>.
- Sharma, S., Sumesh, K.C., Sirisomboon, P., 2022. Rapid ripening stage classification and dry matter prediction of durian pulp using a pushbroom near infrared hyperspectral imaging system. *Measurement* 189. <https://doi.org/10.1016/j.measurement.2021.110464>.
- Somsri, S., 2013. Durian. National Bureau of Agricultural Commodity and Food Standards. [https://www.acfs.go.th/standard/download/DURIAN\\_new.pdf](https://www.acfs.go.th/standard/download/DURIAN_new.pdf).
- Somton, W., Pathaveerat, S., Terdwongworakul, A., 2015. Application of near infrared spectroscopy for indirect evaluation of 'Monthong' durian maturity. *Int. J. Food Prop.* 18 (6), 1155–1168. <https://doi.org/10.1080/10942912.2014.891609>.
- Talabnark, A., Terdwongworakul, A., 2017. Minimally destructive evaluation of durian maturity using near infrared spectroscopy. *Thai Soc. Agric. Eng. J.* 23 (2). [http://www.tsae.asia/new/wp-content/uploads/2018/08/v23\\_2.pdf](http://www.tsae.asia/new/wp-content/uploads/2018/08/v23_2.pdf).
- Timkhun, P., Terdwongworakul, A., 2012. Non-destructive classification of durian maturity of 'Monthong' cultivar by means of visible spectroscopy of the spine. *J. Food Eng.* 112 (4), 263–267. <https://doi.org/10.1016/j.foodeng.2012.05.018>.
- Timkhun, P., Terdwongworakul, A., 2013. Non-destructive classification of durian maturity of 'Monthong' cultivar by visible spectroscopy of the husk. *Thai Society of Agricultural Eng. J.* 19, 1–6. [www.tsae.asia](http://www.tsae.asia).
- Tongdee, S.C., Suwanagul, A., Neamprem, S., 1990. Durian Fruit Ripening and Effect of Variety, Maturity Stage at Harvest, and Atmospheric Gases. *International Society for Horticultural Science*.