IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 17 July 2023, accepted 31 July 2023, date of publication 7 August 2023, date of current version 15 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3303020

APPLIED RESEARCH

Indirect Prediction of Dry Matter in Durian Pulp With Combined Features Using Miniature NIR Spectrophotometer

AMORNRIT PUTTIPIPATKAJORN[1], ANUPUN TERDWONGWORAKUL[2], AMORNDEJ PUTTIPIPATKAJORN[3], SUPACHAI KULMUTIWAT[2], PEERAPONG SANGWANANGKUL[4], AND THANA CHEEPSOMSONG[5]

[1]Department of Computer Engineering, Faculty of Engineering at Kamphaeng Saen, Kasetsart University, Kamphaeng Saen, Nakhon Pathom 73140, Thailand
[2]Department of Agricultural Engineering, Faculty of Engineering at Kamphaeng Saen, Kasetsart University, Kamphaeng Saen, Nakhon Pathom 73140, Thailand
[3]Department of Food Engineering, Faculty of Engineering at Kamphaeng Saen, Kasetsart University, Kamphaeng Saen, Nakhon Pathom 73140, Thailand
[4]Postharvest Technology Center, Department of Horticulture, Faculty of Agriculture at Kamphaeng Saen, Kasetsart University, Kamphaeng Saen, Nakhon Pathom 73140, Thailand
[5]Department of Mechanical Engineering, Faculty of Engineering at Kamphaeng Saen, Kasetsart University, Kamphaeng Saen, Nakhon Pathom 73140, Thailand

Corresponding author: Thana Cheepsomsong (fengtnc@ku.ac.th)

This work was supported by Kasetsart University, Kamphaeng Saen Campus, Thailand, under Grant KPS-RDI2022-008.

ABSTRACT Immature durian poses a crucial problem for durian export. A miniature near-infrared spectrometer was used to develop a prediction model based on fruit stem and rind spectra. A sample of 120 durian fruits with variation in maturity was used to calibrate the model using different machine-learning algorithms, consisting of partial least squares regression (PLSR), least square support vector machine (LS-SVM), and artificial neural network (ANN) approaches. Initially, the rind model provided better predictive performance than the stem model for all algorithms. However, combining either the stem or the rind spectra (or even the fruit density) as the independent variables, the built models produced better results in prediction than from using either the stem or rind spectra alone. Among the models investigated, the combined spectra model of the stem and rind without spike spectra and using the ANN algorithm produced the best prediction with a correlation coefficient of 0.848, a root mean square error of prediction of 2.215%, and a residual predictive deviation of 1.833. Finally, the results showed that the proposed model performed sufficiently well in predicting the dry matter content of durian fruit using the miniature near-infrared spectrometer, even though it underperformed the model derived from the durian pulp.

INDEX TERMS Agricultural engineering, spectroscopy, regression analysis.

I. INTRODUCTION

Durian is an important export fruit of Thailand, with the main durian planting areas being in the east and the south. The quality and good taste of Thai durian has resulted in it being accepted by consumers in many foreign countries. Durian fruits are round-to-oblong in shape and the rind is covered with sharp thorns that are a green-to-yellowish green or brown color (Fig. 1). The stem connects the fruit to the branch and has an abscission zone separating the stem into two parts. Inside each carpel is an edible portion known as the pulp that is golden-yellow in color. The problem with exporting durian during the period before the optimal harvest time has always been the mixing of inferior quality durian, especially immature ones, which, regardless of the cause, has had a negative effect on the country's durian export business. This problem is caused by the mistake of harvesting durians that do not have optimal maturity and is one of the reasons why the value of durian has decreased in price. At present, harvesting mature durian requires experience and expertise. Durian farmers will determine when to harvest based on counting the days after anthesis (DAA), squeezing the ends

The associate editor coordinating the review of this manuscript and approving it for publication was Norbert Herencsar.

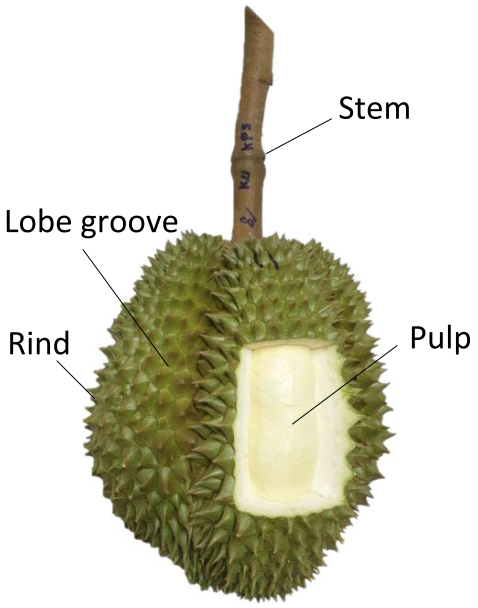


FIGURE 1. Durian fruit showing stem, lobe groove, rind, and pulp.

of the thorns together (those of mature durians require less force to squeeze than those of immature durians), inspecting whether the lobe groove is wide (if it is, the durian is very mature), or tapping while listening (if the sound is clear, the fruit is mature). However, these methods used by farmers and human sorters at the packing house can potentially cause large errors in selection, especially if the person doing the testing is tired or inexperienced [1].

A durian maturity index has been developed for separating immature durian from mature durian by using 32% dry matter content (DMC) as a standard value for identifying the degree of maturity of durian which has been academically accepted [2], [3]. Consequently, the minimum DMC required for mature 'Monthong' durian fruit based on the Thai Agricultural Standard for Durian (TAS 3-2013) is 32%. However, the DMC is the value obtained based on destructive measurement of the durian fruit and it takes time to quantify the result. Therefore, there is a need for a technique that can be used to rapidly measure durian maturity without causing any damage to the durian fruit. Such a technique should be able to repeatedly measure the maturity of fruit based on the DMC, until it is the correct time to harvest.

Near infrared spectroscopy (NIR) is an accepted technique that is rapid and non-destructive for the measurement of the internal composition of agricultural products and food samples. The NIR technique has been applied to accurately assess the quality of thin-skinned fruits such as mangoes [4], strawberries [5], passion fruit [6], and apples [7]. In the case of thick-skinned fruits, an application of the NIR technique was used to predict the total soluble solids (TSS) in melons [8], [9], [10], watermelons [11], and cantaloupes [12]. Other applications included the isolation of melon genotypes [13] and the examination of cracked shells in coconuts [14].

With regard to thick-rind durian, [15] applied an NIR technique to create a model for predicting the DM by measuring the light absorbance in the diffuse reflection pattern of the stem and rind and to study the feasibility of using the light absorption data of the stem and rind to classify durian into 3 durian maturity groups. A multivariate discriminant analysis technique was used to model the light absorption data of the stem, rind, and stem plus rind, producing durian classification accuracy percentages of 91.7%, 84.7%, and 91.4%, respectively. Data from the stalk alone gave greater accuracy in identifying durian than the data from the husk alone. One reason was because the absorption spectra of the stem were close to the absorption spectra of the flesh. However, the highest classification accuracy was achieved by combining the stem and rind data to create the model. In another study, a portable NIR spectrometer at wavelengths in the range 700–1100 nm was used to compare the NIR absorbance measurements for the stem and the rind to create a TSS prediction equation for the pulp [16]. These researchers found that the calibration equation generated from the rind spectrum had a higher level of accuracy than from the stem spectrum, with a correlation coefficient (r) of 0.94 and a standard error of prediction (SEP) of 2.95%.

NIR applications are limited when very little NIR light can penetrate the sample. Thus, with durian, light may not be able to penetrate the thick rind to reach the flesh and be reflected to record a measurement. Reference [17] studied the reduction effect from the rind and developed an empirical technique to adjust the NIR absorption spectra from the whole fruit (NIR peel and pulp) to be close to the NIR absorption spectrum of the flesh, using peel moisture as the adjustment parameter. The equation constructed from the adjusted predicted spectral values to predict the DMC was similar to that generated from the durian flesh spectrum.

The near infrared hyperspectral imaging (NIR-HSI) technique has advantages over NIR because it can measure spatial absorption spectra, thus obtaining more spectral information. As a result, equations generated from hyperspectral data have higher predictive accuracy than equations generated from normal spectra data. NIR-HSI has been applied to measure the DM [18] and had higher accuracy than from using NIR [19]. The NIR-HSI system could classify the ripeness into 3 levels with an accuracy of 93.6% and predict the DMC with a coefficient of determination of prediction (Rp2) higher than 0.80 and a root mean square error of prediction (RMSE) of less than 1.6%.

A review of published research suggested that there has been no application of an NIR miniature spectrometer for DMC prediction in durian. The NIR miniature spectrometer is a portable device with advantages of being low cost and applicable for online sorting using fiber optics. Additionally, few researchers have integrated the NIR absorbance from more than one position on the durian fruit in the development of the prediction model. Thus, the current research aimed to develop a calibration model from rind and stem absorbance values acquired using the miniature spectrometer,

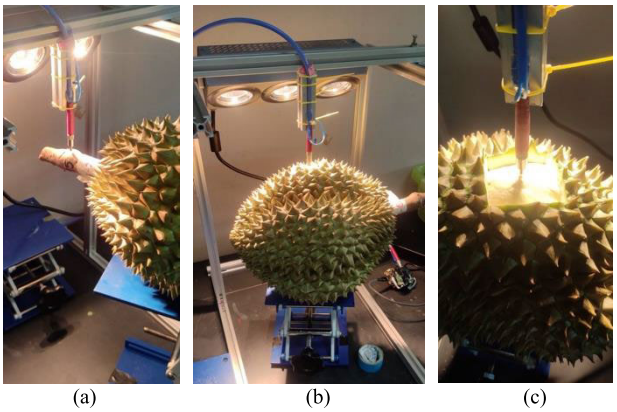


FIGURE 2. Acquisition of NIR reflectance at (a) stem of durian, (b) rind of durian, and (c) durian pulp.

for predicting the DM and to study the integration of the rind and stem absorbance to increase the predictive ability.

II. MATERIALS AND METHODS

A. PREPARATION OF DURIAN SAMPLES

'Monthong' durians were harvested from Chanthaburi province, Thailand. Six harvest ages of durian fruit were randomly picked to cover the range from immature to fully mature fruit, based on DAA with 20 fruits per harvest age: 92 (immature), 99 (premature), 106 (early mature), 113 (mature), 120 (mature), and 127 (fully mature). The harvested fruits were transported to the laboratory at the Postharvest Technology Center, Department of Horticulture, Faculty of Agriculture at Kamphaeng Saen, Kasetsart University, Kamphaeng Saen Campus, Thailand.

B. MEASUREMENT BY VARIOUS TECHNIQUES

1) MEASUREMENT OF NEAR INFRARED SPECTRUM

The absorbance of durian was measured using a miniature spectrometer (DLP Nanoscan; Texas Instruments; USA) in the wavelength range 900–1700 nm, every 3.5 nm in diffuse reflectance measurement mode using fiber optics. Each fruit was measured at 4 locations: (1) at the abscission zone or joint between upper and lower fruit stem (Fig. 2(a)); (2) at the rind surface area between the spikes in the middle of the major lobe (rind) (Fig. 2(b)); (3) at the surface area after one spike had been cut off using knife, which was adjacent to the previously measured surface area (rind area after spike removal) to provide flat surface area for scattering reduction; and (4) at the flesh or aril after rind removal (pulp) (Fig. 2(c)). Absorbance data were acquired by measuring the durian fruit from different ages and recording the absorbance (Fig. 2).

In each measurement, the dark current (I_d) was measured automatically inside the camera, while the white reference (I_w) light intensity was acquired from a spectralon bar with reflective properties in the near-infrared region; these were used to calculate the reflectance (R) according to (1):

$$R = \frac{I_s - I_d}{I_w - I_d} \tag{1}$$

Reflectance data were analyzed using the Matlab software (ver 2021a; MathWorks; MA, USA) to construct dry matter content prediction equations. Savitzky–Golay smoothing (SGS), multiplicative scatter correction (MSC), standard normal variate (SNV), first derivative (1D), and second derivative (2D) values were individually and in combination applied to minimize the light scattering effects. Due to the physical differences of the durian skin on the different fruits used in the analysis, the R, SGS-R, 1D-R, 2D-R, and SNV-R values of the durian fruit rind and pedicle were analyzed separately to create equations. All created models were compared for their DMC predicting performance.

2) DURIAN DENSITY MEASUREMENT

Three-dimensional (3D) scanning [20] was used to determine the volume of each durian fruit. Each fruit was weighed on a digital scale and the durian density was calculated from its weight divided by its volume.

3) DETERMINATION OF DRY MATTER CONTENT

The pulp of the lobe from which the absorbance was measured was used to determine the DMC by chopping the durian pulp thoroughly and obtaining a sample of 20 g by weighing on a digital scale. Then, the sample was dried in a hot-air oven at 70°C for 48 h or until the weight was stable, before calculating the DMC [2]. DMC (%) $= W_1 / W_2 \times 100$, where W_1 is the weight of the sample before drying and W_2 is the weight of the sample after drying.

III. CALIBRATION MODELS

A. PARTIAL LEAST SQUARES REGRESSION (PLSR)

In spectroscopy, a calibration model is generally constructed using the partial least squares (PLS) algorithm, because it performs well in many applications, including the evaluation of fruit quality [21], [22], [23]. The concept behind partial least squares is to decompose both the design matrix X and the response matrix Y [24] by considering information from each other, as demonstrated in (2). This decomposition allows for regression between T and U to be performed.

$$\begin{aligned} X &= TP^T \\ Y &= UQ^T \end{aligned} \tag{2}$$

One intuitive way to achieve this is through an iterative method called the Nonlinear Iterative Partial Least Squares (NIPALS) algorithm, which involves the following procedures.

$u = y_j$ for some j

Loop

$$\begin{aligned} p &= X^T u / \left\| X^T u \right\| \\ t &= Xp \\ q &= Y^T t / \left\| Y^T t \right\| \\ u &= Yq \end{aligned} \tag{3}$$

Until t stop changing

The above provides the procedure for finding the initial set of PLS components and loadings. For subsequent components and vectors, the X and Y are extracted in (4) and used in the same procedure (3) for repetition.

$$X = X - tp^T$$
$$Y = Y - uq^T \tag{4}$$

After repeating the aforementioned steps, matrices T and U, and matrices P and Q, which are related by equation (2), are obtained. A regression model that relates Y and X is achieved by fitting a regression coefficient (β) with U and T, as shown in equation (5).

$$U = T\beta \tag{5}$$

Next, the resulting model is substituted into equation (2) to obtain the calibration model, which predicts the corresponding 'y' value for any given 'x' as follows.

$$Y = UQ^T = T\beta Q^T = XP\beta Q^T \tag{6}$$

In the current study, spectral data acquired from the durian samples and their dry matter contents were considered as the independent and dependent variables, respectively. In the PLS approach adopted, both the spectral data and dry matter content were simultaneously decomposed into orthogonal space called latent variables (LVs). The first several LVs are usually sufficient to maximize the covariance between the X and Y matrices [25] to attain satisfactory calibration and prediction results. The optimal number of LVs was determined using a leave-out-out cross-validation technique. The LVs that resulted in the minimum mean square error (MSE) were considered as the optimal LVs for model construction.

B. LEAST SQUARES SUPPORT VECTOR MACHINE REGRESSION (LS-SVM)

A support vector machine (SVM) is a type of supervised machine-learning classification algorithm that performs very well with even a limited amount of data. SVM constructs a hyperplane in multidimensional space and finds a maximum-marginal hyperplane learned from training data using an optimization procedure to best divide the dataset into classes [26]. It can easily handle multiple continuous and categorical variables, as well as supporting regression problems by modelling the function with a minimum amount of allowable error [27]. The SVM algorithm is implemented using a kernel. A kernel transforms an input data space into the required form, depending on the kernel type, such as a linear, nonlinear, polynomial, radial basis function (RBF) and sigmoid kernel [28]. The RBF kernel function is popular in support vector machine problems and is usually chosen for non-linear data because it provides proper separation when there is no prior knowledge of the data. The SVM algorithm has become very popular in many fields of application, including NIR spectroscopy [29], [30], [31]. SVM for regression is formulated in solving a convex quadratic programming (QP) problem which requires high computation for the constrained optimization. The least squares version of SVM is called LS-SVM [32] and has been proposed to work with equality instead of inequality constraints and a sum squared error cost function. This reformulation with a linear system is known as the Karush–Kuhn–Tucker method for finding the solution and greatly simplifies the problem [33]. The current study selected an approach using the LS-SVM method with an RBF kernel as an alternative algorithm for establishing the calibration model to predict the dry matter in a durian fruit. In LS-SVM regression, the equality constraint-based QP problem is given by (7):

$$\min_{w,b,e_i} J(w, e_i) = \frac{1}{2}w^T w + \frac{1}{2}\gamma \sum_{i=1}^{N} e_i^2$$
$$\text{subject to } y_i = w^T \varphi(x_i) + b + e_i, i = 1, \ldots, N \tag{7}$$

where, w is the weighting vector in the dimension of feature space; b is the bias term; e_i is the error variable; γ is the regularization parameter which penalizes the error and asmaller γ can avoid overfitting noisy data; and $\varphi(.)$ is a kernel space mapping function that maps the data into a higher dimensional feature space. According to the optimization equation, the Lagrange function is obtained as (8):

$$L(w, b, e_i, \alpha_i) = \frac{1}{2}w^T w + \frac{1}{2}\gamma \sum_{i=1}^{N} e_i^2 - \sum_{i=1}^{N} \alpha_i \left\{ w^T \varphi(x_i) + b + e_i - y_i \right\} \tag{8}$$

where, α_i is the set of Lagrange multipliers called support values. The Karush-Kuhn-Tucker conditions for optimality can be obtained by partially differentiating with respect to each variable using:

$$\begin{cases} \frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{i=1}^{N} \alpha_i \varphi(x_i) \\ \frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^{N} \alpha_i = 0 \\ \frac{\partial L}{\partial e_i} = 0 \rightarrow \alpha_i = \gamma e_i, i = 1, .., N \\ \frac{\partial L}{\partial \alpha_i} = 0 \rightarrow w^T \varphi(x_i) + b + e_i - y_i = 0, i = 1, .., N \end{cases} \tag{9}$$

Solving the problem using the Lagrange multipliers eventually yields a set of linear equations, which can be written as (10):

$$\begin{bmatrix} 0 & 1_N^T \\ 1_N & \Omega + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{10}$$

where, 1_N is a N x 1 vector with all entries being 1; $y = [y_1; \ldots; y_N]$ is the vector of reference values; Ω is an N x N matrix with $\Omega_{ij} = K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j)$, which must follow Mercer's condition. Eventually, the LS-SVM regressor

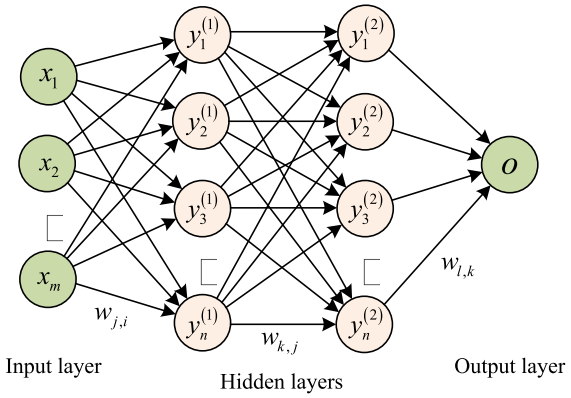


FIGURE 3. Proposed ANN architecture with two hidden layers.

output is expressed as (11):

$$f(x) = \sum_{i=1}^{N} \alpha_i K(x, x_i) + b \tag{11}$$

where, α and b are the solution to (5); $K(x, x_i)$ represents the kernel function that in the current study was a Gaussian radial basis function (RBF) as shown in (12):

$$K(x, x_i) = exp\left(-\|x - x_i\|^2 \Big/ 2\sigma^2\right) \tag{12}$$

In the software implementation, a grid-search technique was used to tune the regularization (γ) and the RBF kernel function parameter (σ^2) in the LS-SVM lab toolbox [34] and the Matlab R2021a software.

C. ARTIFICIAL NEURAL NETWORK (ANN)

An ANN is a subset of machine learning which attempts to simplify and simulate the behavior of the biological systems of a human brain. ANN can process data and solve problems more effectively than conventional machine-learning algorithms. In large-scale datasets with a lot of noise and distorted patterns, it can find interesting features and handle large-dimensional data spaces with high processing speed and robustness [35]. The architecture of ANN involves the arrangement of a network along with its nodes and connecting lines distributed across an input layer, one or more hidden layers, and an output layer. ANN connects nodes in each layer in a forward direction from the input nodes to the hidden nodes to the output nodes with no backward links permitted [36], [37]. In supervised learning, ANN is trained with samples of matched input and output data to generate predicted outputs that are as close as possible.

In the current experiments, the proposed ANN architecture, shown in Fig. 3, had three layers: an input layer, two hidden layers, and an output layer. The linkage lines between adjacent layers presented the network weights: $w_{(j,i)}$ (between the input and the 1st hidden layer), $w_{(k,j)}$ (between the 1st and 2nd hidden layer) and $w_{(l,k)}$ (between the 2nd hidden layer and the output layer), respectively. In the input layer, the number of nodes was determined by the number of spectral wavelengths of the samples. We restricted the number of hidden layers to two to prevent overfitting and to regulate the model's complexity. This decision was made to avoid creating a model that would be overly intricate relative to the relatively small input dataset that consisted of only 85 training samples and 35 validation samples. The number of neural units in each layer was adjusted from 1 to 30 to find an optimal solution with the lowest possible error (MSE).

The output for each node in the network can be calculated in the forward direction using (13):

$$\begin{aligned} y_j^{(h)} &= f\left(w_{j,1}^{(h)} y_1^{(h-1)} + w_{j,2}^{(h)} y_2^{(h-1)} + \cdots + w_{j,m}^{(h)} y_m^{(h-1)} + b_j^{(h)}\right) \\ &= f\left(\sum_i^m w_{j,i}^{(h)} y_i^{(h-1)} + b_j^{(h)}\right) \end{aligned} \tag{13}$$

where $y_i^{(0)} = x_i$ and $y_j^{(N)} = o$ and N is the number of layers in the network, m is the number of inputs, i indicates the current node in the previous layer, and j and h are the current node and the current layer, respectively, $f(\cdot)$ is the transfer function, using a tan-sigmoid transfer function.

During a training process, a neural network's parameters can be adjusted using Levenberg–Marquardt optimization (LM) [38]. Typically, the LM algorithm is the fastest backpropagation algorithm compared to other learning algorithms. It combines two numerical minimization algorithms (the gradient descent method and the Gauss–Newton method). The gradient descent method updates the parameters in the direction of the steepest descent to minimize errors. The Gauss–Newton method minimizes errors differently by assuming that the least-squares function is locally quadratic in the parameters and finding the minimum of this quadratic. When the parameters are far from their optimal value, the LM method behaves more like a gradient-descent technique. When the parameters are nearly at their optimal value, it behaves more like the Gauss–Newton method.

The LM algorithm was developed especially for loss functions. It works with the gradient vector and the Jacobian matrix instead of calculating the exact Hessian matrix. The loss function takes the form of a sum of squared errors (SSE), as shown in (14):

$$f = \sum_{i=1}^{m} e_i^2 \tag{14}$$

where f is the loss function, e is the sum of squared errors, and m is the number of training samples in the data set. We can define the Jacobian matrix of the loss function as the derivatives of the errors for the network parameters using (15):

$$\boldsymbol{J}_{i,j} = \frac{\partial e_i}{\partial \boldsymbol{w}_j}, \tag{15}$$

for $i = 1, \ldots, m$ and $j = 1, \ldots, n$, where n is the number of parameters in the neural network. Note that the size of the Jacobian matrix is $m \times n$. We can compute the gradient vector of the loss function using (16):

$$\nabla f = 2\boldsymbol{J}^T \cdot \boldsymbol{e} \tag{16}$$

Lastly, we can approximate the Hessian matrix using (17):

$$\boldsymbol{H}f \approx 2\boldsymbol{J}^T \cdot \boldsymbol{J} + \lambda \boldsymbol{I} \tag{17}$$

where λ is the damping factor that ensures the positiveness of the Hessian and $\boldsymbol{I}$ is the identity matrix. Equation 18 defines the parameter updating process with the LM algorithm:

$$\boldsymbol{w}^{(i+1)} = \boldsymbol{w}^{(i)} - \left(\boldsymbol{J}^{(i)T} \cdot \boldsymbol{J}^{(i)} + \lambda^{(i)}\boldsymbol{I}\right)^{-1} \cdot \left(2\boldsymbol{J}^{(i)T} \cdot \boldsymbol{e}^{(i)}\right) \tag{18}$$

If the damping parameter λ is zero, then this is simplified to Newton's method using an approximate Hessian matrix. On the other hand, when λ is large, this turns into a gradient descent with a low training rate. The parameter λ should initially be large for the initial updates to be small steps in the gradient descent direction. If any iteration is unsuccessful, then λ is increased by some factor. Otherwise, as λ is reduced, the loss decreases and the LM algorithm gets closer to the Newton method. Typically, convergence is accelerated to the minimum through this process.

As noted previously, the LM algorithm is a technique designed specifically for the sum of squared error functions and this speeds up neural network training. However, this algorithm has some shortcomings. First, it is unable to minimize error functions, such as the cross-entropy error or the root mean squared error. In addition, the Jacobian matrix calculation consumes a lot of memory. Therefore, the LM algorithm is not suggested for large datasets [39].

IV. EXPERIMENTAL DESIGN AND MODEL DEVELOPMENT

In the model development phase, the collection of the 120 durian fruits was divided into two groups using a 70:30 ratio, consisting of a calibration set of 85 and a prediction set of 35, respectively. Each group was made up of durian fruit harvested at the different ages, ranging from immature to fully mature, based on DAA. Due to limitations in the collection of samples, most samples were within the range of DAA 106 (DMC $\approx$ 33%) to DAA 127 (DMC $\approx$ 40%), while only a few samples were collected between DAA 92–99, with a DMC less than 33%. However, the selection of fruit was randomized and the range covered DAA 92 (immature) to DAA 127 (fully mature). The calibration set was arranged for each given schema before being fed as an independent variable (input dataset) into the model. The independent variables used to construct the calibration models consisted of the NIR reflectance spectra of the durian stem (SR), NIR reflectance spectra of the durian rind (RR), NIR reflectance spectra of the durian rind without spikes (RNSR), NIR reflectance spectra of the durian pulp (PR), and the density of the durian (DEN), respectively. The DMC of the durian was used as a dependent variable to construct the model. Initially, the DMC of all samples was arranged in ascending order before building the model. The first and last samples were chosen for the calibration group, while the remaining samples were alternatively chosen for the calibration group and the validation group. As a result, the samples from both groups had similar DMC distributions, with the DMC of the samples from the calibration group covering the DMC of the samples from the validation group.

TABLE 1. Changes in durian physical properties with DAA.

DAA	Average fruit mass (g)	Average fruit volume (cm^3)	Average fruit density (g/cm^3)	Average dry matter content (%)
85	3764.8	4371.0	0.862	27.36
92	3639.5	4234.4	0.861	33.33
99	3695.8	4332.5	0.851	35.97
106	3738.8	4367.7	0.856	36.52
113	3294.0	3968.6	0.830	36.64
120–127	2701.8	3284.3	0.823	38.63

DAA: Days after anthesis

The regression models presented in this paper (PLS, LS-SVM, and ANN) were specifically designed for a regression problem investigating the dry matter content of durian fruit. This paper presented some commonly used regression metrics, consisting of the correlation coefficient (r), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE), as shown in (19)– (21), respectively:

$$r = \sqrt{1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}} \tag{19}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{n}} \tag{20}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|y_i - \hat{y}_i\right|}{|y_i|} \times 100\% \tag{21}$$

where y_i is the reference value, $\bar{y}$ is the mean of y, $\hat{y}_i$ is the predicted value of y_i, and n is the number of samples.These regression metrics are statistical parameters and helped to assess how well each regression model predicted the actual values of the target variable. In addition, these metrics helped in the selection of the best regression model among the various tested models and to optimize the hyperparameters of the model for better performance.

In the current study, the calibration models of the various algorithms (PLSR, LS-SVM, ANN) with different spectral pre-treatments (SGS, SNV, MSC, 1D, 2D) were constructed using different kinds of input datasets. To find the best predictive results, three schemas were proposed, as shown in Fig. 4. In the first schema, individual NIR spectra (SR, RR, RNSR) were used as the input dataset for constructing the models and produced 15 different models (an input dataset generates 5 models using different spectral pre-treatments for each algorithm), whereas the second and third schemas presented the combination of input dataset of NIR spectra (SR+RR and SR+RNSR) and the combination of input dataset of NIR spectra and the fruit density (SR+DEN, RR+DEN and RNSR+DEN), respectively. Each combined input dataset produced 10 and 15 models that were expected to improve the prediction results compared with individual input datasets.

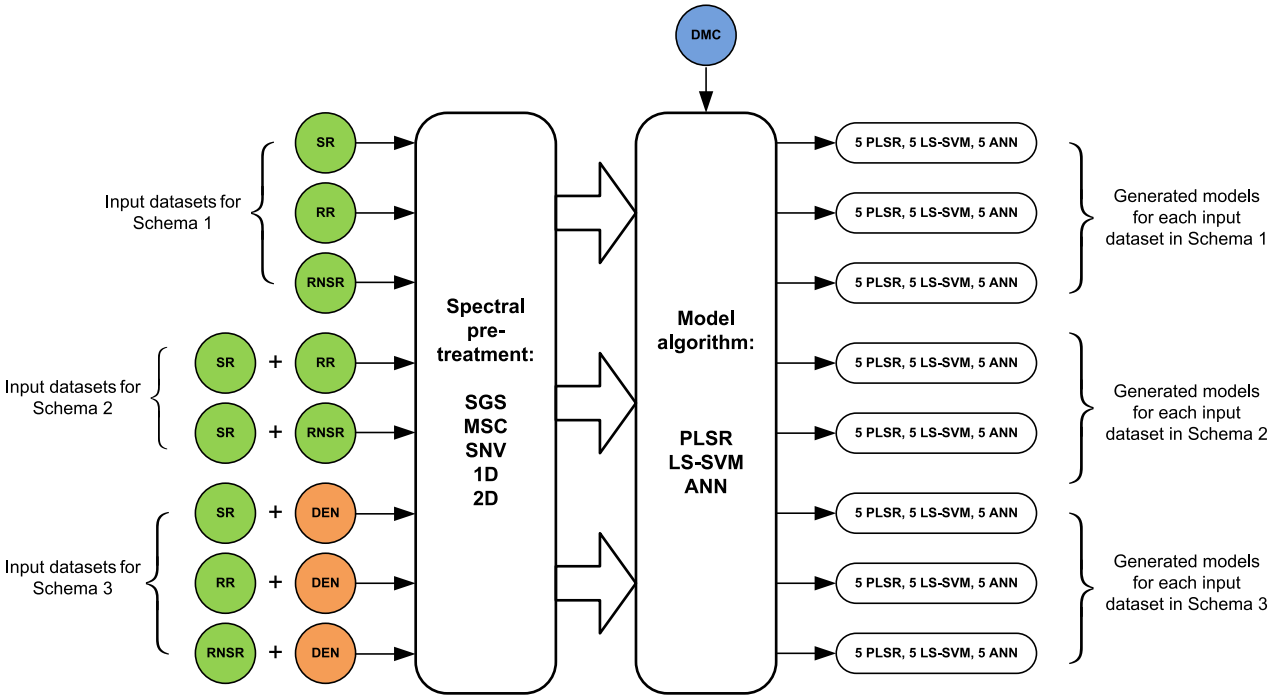


FIGURE 4. Schemas for calibration model construction using different input datasets.

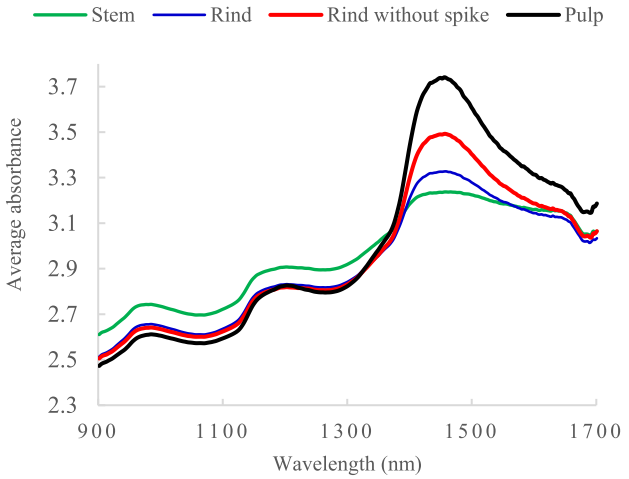


FIGURE 5. Comparison of NIR spectra of stem, rind, rind without spike, and pulp.

V. RESULTS AND DISCUSSION

A. CHANGE IN DURIAN DRY MATTER CONTENT

Monthong durian from Chanthaburi province, Thailand (randomly picked from 6 age groups after flowering, with 20 fruits per age group) had weight, volume, and density values, as shown in Table 1. The average fruit density of each age tended to decrease with increasing age after flowering or maturity, while the dry matter content increased, respectively, with age after flowering. This pattern of dry matter content gain was consistent with other research [1], [2], [15].

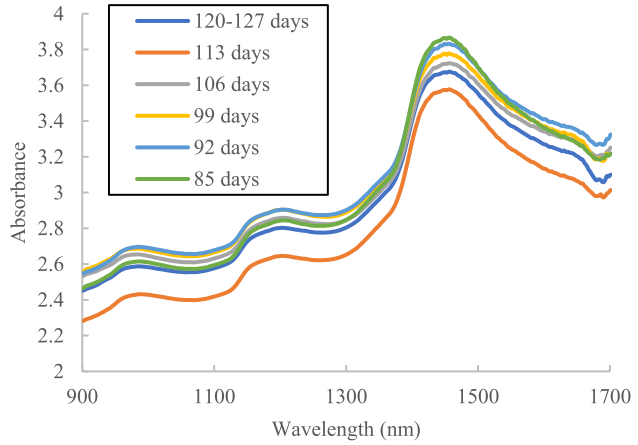
B. SPECTRAL CHARACTERISTICS

The NIR absorption spectra of the stem, rind, thornless rind and flesh are compared in Fig. 5. Dominant peaks were observed at around 970, 1200, and 1450 nm. The peaks at around 970 and 1200 nm were assigned to the absorption bands of the water molecule vibrating in the symmetric and asymmetric stretching modes [40]. A major peak at around 1450 nm, where the pulp absorbance was higher than for the stem and the rind, was associated with the O-H stretching first overtone of starch and water [41], indicating that the fruit pulp contained more moisture than the stem and peel. The stem had the lowest moisture content. Overall, the absorbance of the shell other than the water absorbance region was lower than for, which was consistent with the findings of [15]. This was due to the peel having durian thorns, causing the light to scatter more than for the smoother stalk, resulting in the shell reflecting more light than the stalk or absorbing less light than the stalk itself.

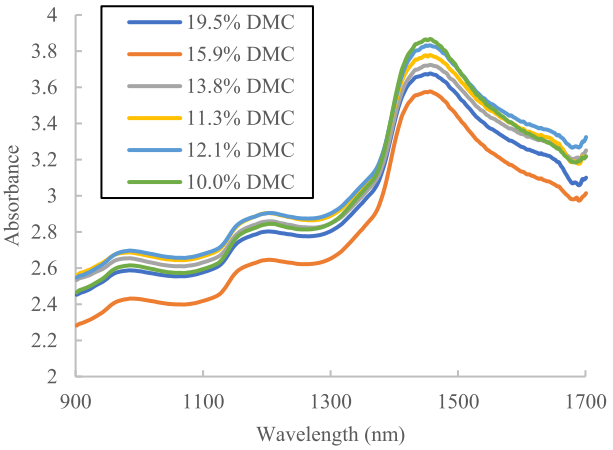
The effect of DAA on the absorbance of flesh is shown in Fig. 6(a), based on the mean absorbance of durian flesh in each maturity stage. If the absorbance is divided into 3 groups, where DAA from 85-92 days is the immature group, 99-106 days is the early mature group, and 113-127 days is the mature group, it can be seen that the absorbance around the wavelengths from 1400 to 1600 nm tended to decrease with maturity, indicating that the pulp absorbed less light as the fruit became more mature. The variation in pulp absorbance concerning DMC is alternatively displayed in Fig. 6(b). The absorbance around the 1450 nm wavelength, which corresponds to a water peak, exhibited

TABLE 2. Best predictive performance of different calibration models based on Individual NIR spectra from various locations according to Schema 1.

Algorithm	Input dataset	Pre-treatment	LVs/hidden nodes	Rc	RMSEc	MAPEc	Rp	RMSEp	RPDp	MAPEp
PLS	SR	MSC	6	0.742	3.349	8.284	0.607	3.180	1.277	7.594
	RR	1D	6	0.730	3.417	8.246	0.668	2.978	1.364	7.163
	RNSR	1D	12	0.861	2.543	5.930	0.721	2.773	1.464	6.992
	PR	**SNV**	**12**	**0.917**	**1.991**	**4.800**	**0.882**	**1.887**	**2.152**	**4.162**
LS-SVM	SR	MSC	-	0.927	1.873	4.381	0.706	2.835	1.432	6.620
	RR	1D	-	0.731	3.412	5.080	0.643	3.066	1.324	7.295
	RNSR	1D	-	0.853	2.610	6.006	0.738	2.700	1.504	6.361
	PR	**SNV**	**-**	**0.949**	**1.571**	**3.483**	**0.888**	**1.839**	**2.208**	**4.180**
ANN	SR	SGS	3, 14	0.945	1.685	3.555	0.773	2.568	1.581	6.089
	RR	SGS	7, 4	0.877	2.430	5.815	0.783	2.513	1.616	5.969
	RNSR	1D	20, 3	0.901	2.169	4.716	0.818	2.358	1.722	5.588
	PR	**MSC**	**27, 6**	**0.968**	**1.346**	**3.001**	**0.919**	**1.706**	**2.380**	**4.129**



a)



b)

FIGURE 6. Change in absorbance of pulp with respect to (a) days after anthesis (DAA) and (b) dry matter content (DMC).

a decreasing trend as DMC increased. This trend was consistent with other studies [15], [18]. However, over time, the absorption tendencies of the stem, rind, and the rind without spikes were not as clear as for the flesh (data not shown).

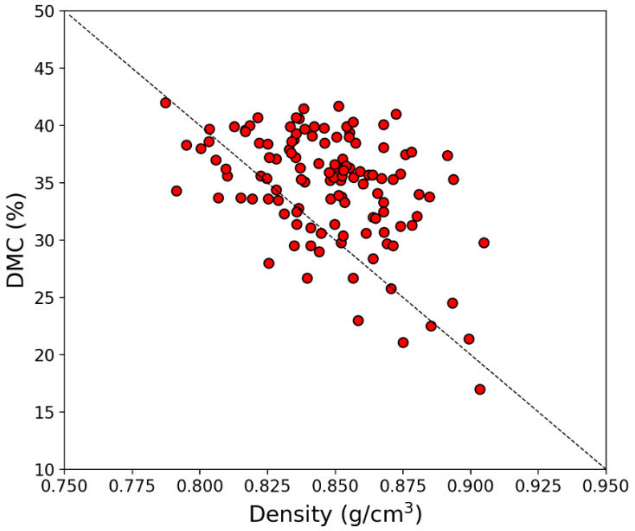


FIGURE 7. Relationship between density of durian fruit and dry matter.

C. RESULTS OF DATA ANALYSIS FOR PREDICTING DRY MATTER CONTENT

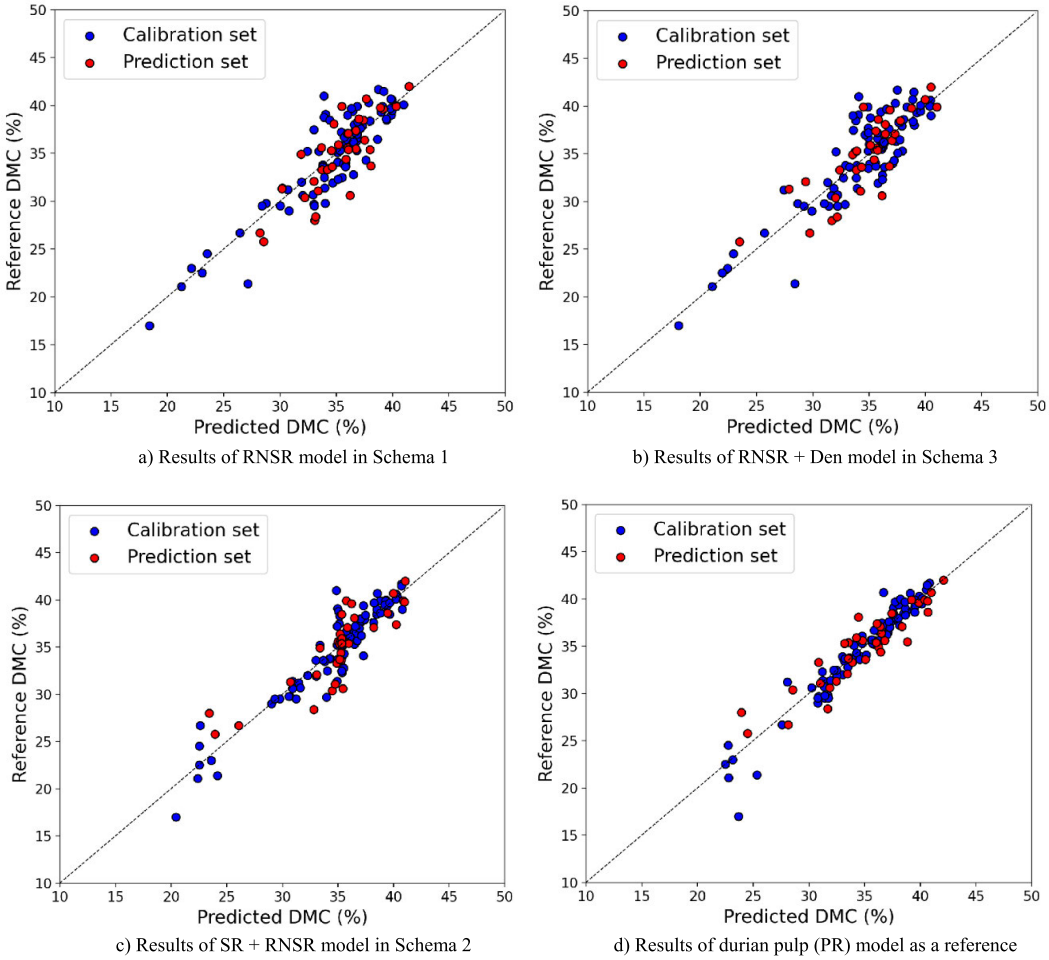
The NIR spectra for the three different locations on the surface of the durian (SR, RR, and RNSR) were used to create regression models (PLSR, LS-SVM, ANN) to evaluate the DMC of the durian. The NIR spectral data of the durian pulp (PR) were not included in the development of the model. Instead, these data were used to construct a separate model that was compared to the models generated from the NIR spectra for the surface of the durian. Considering the generated models according to Schema 1, Table 2 presents the best predictive outcomes from the different models (PLSR, LS-SVM, ANN) for the different fruit measurement locations. The models of RNSR provided more accurate results than the models of RR and RNSR with values for RMSE of 2.773 for PLSR, of 2.700 for LS-SVM, and of 2.358 for ANN; the ANN algorithm clearly outperformed the PLSR and LS-SVM ones. These results matched the work of [15] that showed that the models derived from the rind spectrum offered a more precise prediction of the dry matter content of durian than the models derived from the stem spectrum. In addition, this finding confirmed

TABLE 3. Best predictive performance of different calibration models based on combined NIR spectra according to Schema 2.

Algorithm	Input dataset	Pre-treatment	LVs/Hidden nodes	Rc	RMSEc	MAPEc	Rp	RMSEp	RPDp	MAPEp
PLSR	**SR + RR**	**1D**	**7**	**0.796**	**3.024**	**7.239**	**0.709**	**2.821**	**1.439**	**7.001**
	SR + RNSR	SGS	7	0.786	3.091	7.689	0.667	2.981	1.362	7.133
LS-SVM	SR + RR	SNV	-	0.773	3.171	7.532	0.662	3.00	1.354	7.075
	SR + RNSR	**1D**	**-**	**0.942**	**1.682**	**5.283**	**0.686**	**2.911**	**1.395**	**6.799**
ANN	SR + RR	SNV	3, 14	0.881	2.392	5.465	0.783	2.545	1.595	5.974
	SR + RNSR	**1D**	**15, 2**	**0.940**	**1.708**	**3.639**	**0.848**	**2.215**	**1.833**	**5.178**

TABLE 4. Best predictive performance of different calibration models based on combination of NIR spectra and fruit density according to Schema 3.

Algorithm	Input dataset	Pre-treatment	LVs/Hidden nodes	Rc	RMSEc	MAPEc	Rp	RMSEp	RPDp	MAPEp
PLS	SR + DEN	SNV	6	0.748	3.316	8.264	0.642	3.07	1.323	7.571
	RR + DEN	1D	6	0.726	3.439	8.241	0.685	2.915	1.393	6.830
	RNSR + DEN	**SGS**	**11**	**0.856**	**2.583**	**5.779**	**0.714**	**2.802**	**1.449**	**6.779**
LS-SVM	**SR + DEN**	**SNV**	**-**	**0.998**	**0.247**	**0.457**	**0.758**	**2.608**	**1.557**	**6.258**
	RR + DEN	SGS	-	0.813	2.912	6.891	0.654	3.026	1.342	7.340
	RNSR + DEN	SGS	-	0.862	2.534	5.492	0.722	2.767	1.468	6.524
ANN	SR + DEN	MSC	25, 3	0.948	1.740	4.182	0.807	2.432	1.670	5.897
	RR + DEN	SGS	9, 2	0.877	2.406	5.372	0.787	2.498	1.626	5.776
	RNSR + DEN	**SGS**	**2, 10**	**0.879**	**2.405**	**5.306**	**0.826**	**2.262**	**1.795**	**5.293**



a) Results of RNSR model in Schema 1

b) Results of RNSR + Den model in Schema 3

c) Results of SR + RNSR model in Schema 2

d) Results of durian pulp (PR) model as a reference

FIGURE 8. Scatter plots between reference and predicted values of ANN models that produced best results in each schema.

that the moisture content of the rinds continued to decline as the durian grew older [42].

In terms of spectral pre-treatment, the experiments revealed that the SNV and MSC pre-treatment methods, including SGS, performed well on smooth surfaces like the durian stem and the durian rind without spikes, whereas the first derivative pre-treatment performed better on an uneven surface like the durian rind, which is more susceptible

to light scattering. However, greater complexity in the spectrum derived from derivative processing could include noninformative parts of the spectrum which could deteriorate the model's prediction.

Table 3 presents the best predictive results of the models based on Schema 2, based on two spectrum combinations. Compared to the outcomes of the individual models in Table 2, the combined input models of SR+RR and SR+RNSR showed no appreciable improvement in either the PLS or the LS-SVM algorithms. Their predictive performances were no better than those of the RNSR models. However, for only the ANN algorithm, the combined input model of SR+RNSR, performed notably well, with an RMSE of 2.215 that outperformed the ones of SR (2.568) and RR (2.513). In addition, the results showed that the combined NIR spectra model did not improve the predictive results for PLSR and LS-SVM; however, there was some improvement in the ANN algorithm with the first derivative pre-treatment. This may have resulted from the algorithm's learning process, as well as the more complicated properties of the first derivative spectrum.

Schema 3 produced the best predictive results based on the combination of NIR spectra and fruit density (Table 4). The combined input models of SR+DEN and RR+DEN showed a slight improvement over the models in Table 2, whereas the RNSR+DEN model showed a significant improvement only for the ANN algorithm. The RNSR+DEN model produced the best predictive results compared to the SR+DEN and RR+DEN models. These results showed that fruit density, which has been recognized as a reliable indicator of fruit and vegetable DMC, might help improve the prediction performance of the models [43], [44]. As shown in Fig. 7, the density of durians was also correlated with their dry matter. However, compared to the combined NIR spectra models (SR+RR and SR+RNSR), the RNSR+DEN model with the ANN algorithm (RMSE = 2.262) underperformed the SR+RNSR model (RMSE = 2.215) in predicting the dry matter content.

Based on the overall results, the ANN algorithm outperformed the other calibration models in every test. The predictive results associated with the spectrum of RNSR were better, either as a single or combined feature. This implied that the spectrum at this location contained more informative data compared to the other locations on the fruit. Furthermore, the combination of NIR spectra and fruit density (Schema 3) and the combined NIR spectra at different measuring locations (Schema 2) had an advantage over the individual NIR spectra. There were some improvements in Schema 2 over Schema 3, as depicted in Fig. 8, which showed the best results in terms of the scatter plots for each schema, including the ANN result of the spectrum acquired from the durian pulp used as a comparison.

VI. CONCLUSION

A miniature NIR spectrometer could be utilized to indirectly predict the dry matter content of durian pulp by analyzing the stem spectra, the rind spectra, and the rind without spike spectra. For this purpose, PLS, LS-SVM, and ANN models were developed and compared in terms of their prediction performance for dry matter content. The ANN models outperformed the PLS and LS-SVM models, with the latter two performing equally well. Notably, the model using the rind without spikes exhibited superior performance compared to the models using only the stem and rind spectra. Combining the rind and stem spectra resulted in improved predictions compared to using individual spectra.

Furthermore, the inclusion of fruit density as an additional independent variable in the input dataset, for either the rind or stem spectra, or even for the rind without spikes spectra, enhanced the prediction performance. Based on the RMSE values, the proposed models, especially the combined NIR model based on the stem and the rind without the spike spectra using the ANN algorithm, demonstrated the most acceptable results in indirectly predicting the dry matter content of durian pulp.

However, a limitation of this technique was the need to carry out each measurement separately, leading to a comparatively long time for analysis. To improve the accuracy of the models, future work could focus on increasing the NIR light power, which might provide better spectral information of the rind and stem. Moreover, further research with a larger sample size is necessary to validate the results and enhance confidence in the predictive performance of the model.

ACKNOWLEDGMENT

The Kasetsart University Research and Development Institute (KURDI), Bangkok, Thailand, provided assistance with English editing.

REFERENCES

[1] S. Ketsa, A. Wisutiamonkul, Y. Palapol, and R. E. Paull, "The durian: Botany, horticulture, and utilization," in *Horticultural Reviews*, vol. 47. Hoboken, NJ, USA: Wiley, 2020, pp. 125–211.

[2] P. Sangwanangkul and J. Siriphanich, "Growth and maturation of durian fruit cv. Monthong," *Thai J. Agric. Sci.*, vol. 3, nos. 1–2, pp. 75–82, 2000.

[3] J. Siriphanich, "Durian (*Durio zibethinus* Merr.)," in *Postharvest Biology and Technology of Tropical and Subtropical Fruits*, E. M. Yahia, ed. Sawston, U.K.: Woodhead, 2011, pp. 80e–116e.

[4] S. N. Jha, P. Jaiswal, K. Narsaiah, M. Gupta, R. Bhardwaj, and A. K. Singh, "Non-destructive prediction of sweetness of intact mango using near infrared spectroscopy," *Scientia Horticulturae*, vol. 138, pp. 171–175, May 2012.

[5] M. L. Amodio, F. Ceglie, M. M. A. Chaudhry, F. Piazzolla, and G. Colelli, "Potential of NIR spectroscopy for predicting internal quality and discriminating among strawberry fruits from different production systems," *Postharvest Biol. Technol.*, vol. 125, pp. 112–121, Mar. 2017, doi: 10.1016/j.postharvbio.2016.11.013.

[6] G. A. de Oliveira, F. de Castilhos, C. M.-G.-C. Renard, and S. Bureau, "Comparison of NIR and MIR spectroscopic methods for determination of individual sugars, organic acids and carotenoids in passion fruit," *Food Res. Int.*, vol. 60, pp. 154–162, Jun. 2014.

[7] S. N. Jha and R. Garg, "Non-destructive prediction of quality of intact apple using near infrared spectroscopy," *J. Food Sci. Technol.*, vol. 47, no. 2, pp. 207–213, Mar. 2010.

[8] S. Xu, H. Lu, X. Wang, C. M. Ference, X. Liang, and G. Qiu, "Nondestructive detection of internal flavor in 'Shatian' pomelo fruit based on visible/near infrared spectroscopy," *HortScience*, vol. 56, no. 11, pp. 1325–1330, Nov. 2021.

[9] J. A. Guthrie, C. J. Liebenberg, and K. B. Walsh, "NIR model development and robustness in prediction of melon fruit total soluble solids," *Aust. J. Agric. Res.*, vol. 57, no. 4, pp. 411–418, Apr. 2006.
[10] C. V. Greensill, P. J. Wolfs, C. H. Spiegelman, and K. B. Walsh, "Calibration transfer between PDA-based NIR spectrometers in the NIR assessment of melon soluble solids content," *Appl. Spectrosc.*, vol. 55, no. 5, pp. 647–653, May 2001.
[11] A. A. A. Rahim, H. Hashim, N. E. Abdullah, S. L. M. Hassan, I. Shairah, A. Halim, and F. A. Igol, "A numerical analysis of correlation between sucrose level measurement and near-infrared (NIR) for various grades of watermelon ripeness," in *Proc. Int. Conf. Technol., Informat., Manage., Eng. Environ.*, Jun. 2013, pp. 180–185.
[12] G. G. Dull, G. S. Birth, D. A. Smittle, and R. G. Leffler, "Near infrared analysis of soluble solids in intact cantaloupe," *J. Food Sci., Off. Publication Inst. Food Technol. USA*, vol. 54, no. 2, pp. 393–395, Mar. 1989.
[13] Z. Seregély, T. Deák, and G. D. Bisztray, "Distinguishing melon genotypes using NIR spectroscopy," *Chemometric Intell. Lab. Syst.*, vol. 72, no. 2, pp. 195–203, Jul. 2004.
[14] S. Noypitak, W. Imsabai, W. Noknoi, S. Karoojee, A. Terdwongworakul, and H. Kobori, "Detection of cracked shell in intact aromatic young coconut using near infrared spectroscopy and acoustic response methods," *J. Food Meas. Characterization*, vol. 13, no. 3, pp. 1991–1999, Sep. 2019.
[15] W. Somton, S. Pathaveerat, and A. Terdwongworakul, "Application of near infrared spectroscopy for indirect evaluation of 'Monthon' durian maturity," *Int. J. Food Properties*, vol. 18, no. 6, pp. 1155–1168, Jun. 2015.
[16] N. Tangjitwiboonkun, P. Chaiprasart, and R. Rittiron, "Nondestructive determination of total soluble solids in 'Long-lab-lae' durian fruits using near-infrared spectroscopy," *Acta Horticulturae*, no. 1186, pp. 151–156, Nov. 2017.
[17] K. Phuangsombut, A. Phuangsombut, A. Talabnark, and A. Terdwongworakul, "Empirical reduction of rind effect on rind and flesh absorbance for evaluation of durian maturity using near infrared spectroscopy," *Postharvest Biol. Technol.*, vol. 142, pp. 55–59, Aug. 2018.
[18] S. Sharma, K. C. Sumesh, and P. Sirisomboon, "Rapid ripening stage classification and dry matter prediction of durian pulp using a pushbroom near infrared hyperspectral imaging system," *Measurement*, vol. 189, Feb. 2022, Art. no. 110464.
[19] P. Onsawai and P. Sirisomboon, "Determination of dry matter and soluble solids of durian pulp using diffuse reflectance near infrared spectroscopy," *J. Near Infr. Spectrosc.*, vol. 23, no. 3, pp. 167–179, Jun. 2015.
[20] T. Cheepsomsong and J. Siriphanich, "Durian volume determination using short-range coded-light three-dimensional scanner," *Agric. Nat. Resour.*, vol. 56, no. 1, pp. 112–120, Jan./Feb. 2022.
[21] P. D. Alamar, E. T. S. Caramês, R. J. Poppi, and J. A. L. Pallone, "Quality evaluation of frozen guava and yellow passion fruit pulps by NIR spectroscopy and chemometrics," *Food Res. Int.*, vol. 85, pp. 209–214, Jul. 2016, doi: 10.1016/j.foodres.2016.04.027.
[22] P. Maniwara, K. Nakano, S. Ohashi, D. Boonyakiat, P. Seehanam, P. Theanjumpol, and P. Poonlarp, "Evaluation of NIRS as non-destructive test to evaluate quality traits of purple passion fruit," *Scientia Horticulturae*, vol. 257, Nov. 2019, Art. no. 108712.
[23] M. Wang, Y. Xu, Y. Yang, B. Mu, M. A. Nikitina, and X. Xiao, "Vis/NIR optical biosensors applications for fruit monitoring," *Biosensors Bioelectron.*, vol. 11, Sep. 2022, Art. no. 100197.
[24] H. N. Moghaddam, Z. Tamiji, M. A. Lakeh, M. R. Khoshayand, and M. H. Mahmoodi, "Multivariate analysis of food fraud: A review of NIR based instruments in tandem with chemometrics," *J. Food Composition Anal.*, vol. 107, Apr. 2022, Art. no. 104343.
[25] S. Wold, M. Sjöström, and L. Eriksson, "PLS-regression: A basic tool of chemometrics," *Chemometric Intell. Lab. Syst.*, vol. 58, no. 2, pp. 109–130, Oct. 2001.
[26] M. A. Munnaf and A. M. Mouazen, "Removal of external influences from on-line vis-NIR spectra for predicting soil organic carbon using machine learning," *Catena*, vol. 211, Apr. 2022, Art. no. 106015.
[27] S. Dharumarajan, M. Lalitha, C. Gomez, R. Vasundhara, B. Kalaiselvi, and R. Hegde, "Prediction of soil hydraulic properties using VIS-NIR spectral data in semi-arid region of Northern Karnataka Plateau," *Geoderma Regional*, vol. 28, Mar. 2022, Art. no. e00475.
[28] F. B. de Santana, S. K. Otani, A. M. de Souza, and R. J. Poppi, "Comparison of PLS and SVM models for soil organic matter and particle size using vis-NIR spectral libraries," *Geoderma Regional*, vol. 27, Dec. 2021, Art. no. e00436.
[29] M. L. da Silva Medeiros, Y. L. Brasil, L. J. P. Cruz-Tirado, A. F. Lima, H. T. Godoy, and D. F. Barbin, "Portable NIR spectrometer and chemometric tools for predicting quality attributes and adulteration levels in butteroil," *Food Control*, vol. 144, Feb. 2023, Art. no. 109349, doi: 10.1016/j.foodcont.2022.109349.
[30] Devianti, Sufardi, R. Bulan, and A. Sitorus, "Vis-NIR spectra combined with machine learning for predicting soil nutrients in cropland from Aceh Province, Indonesia," *Case Stud. Chem. Environ. Eng.*, vol. 6, Dec. 2022, Art. no. 100268.
[31] S. Liu, S. Wang, C. Hu, X. Qin, J. Wang, and D. Kong, "Development of a new NIR-machine learning approach for simultaneous detection of diesel various properties," *Measurement*, vol. 187, Jan. 2022, Art. no. 110293.
[32] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Process. Lett.*, vol. 9, no. 3, pp. 293–300, Jun. 1999.
[33] F. Chauchard, R. Cogdill, S. Roussel, J. M. Roger, and V. Bellon-Maurel, "Application of LS-SVM to non-linear phenomena in NIR spectroscopy: Development of a robust and portable sensor for acidity prediction in grapes," *Chemometric Intell. Lab. Syst.*, vol. 71, no. 2, pp. 141–150, May 2004, doi: 10.1016/j.chemolab.2004.01.003.
[34] K. de Brabanter, P. Karsmakers, F. Ojeda, C. Alzate, J. de Brabanter, K. Pelckmans, B. de Moor, J. Vandewalle, and J. A. K. Suykens, "LS-SVMlab toolbox user's guide: Version 1.8," Dept. Elect. Eng., Katholieke Universiteit Leuven, Leuven, Belgium, Tech. Rep. ESAT-SISTA-10-146, 2011.
[35] D. Kruzlicova, J. Mocak, B. Balla, J. Petka, M. Farkova, and J. Havel, "Classification of Slovak white wines using artificial neural networks and discriminant techniques," *Food Chem.*, vol. 112, no. 4, pp. 1046–1052, Feb. 2009.
[36] R. Miikkulainen, "Topology of a neural network," in *Encyclopedia of Machine Learning*, C. Sammut and G. I. Webb, eds. Boston, MA, USA: Springer, 2011, doi: 10.1007/978-0-387-30164-8_837.
[37] D. Svozil, V. Kvasnicka, and J. Pospichal, "Introduction to multi-layer feed-forward neural networks," *Chemometric Intell. Lab. Syst.*, vol. 39, no. 1, pp. 43–62, Nov. 1997.
[38] H. P. Gavin. *The Levenberg–Marquardt Algorithm for Nonlinear Least Squares Curve-Fitting Problems*. [Online]. Available: https://people.duke.edu/ hpgavin/ExperimentalSystems/lm.pdf
[39] D. Pérez-Marín, A. Garrido-Varo, J. E. Guerrero, and J. C. Gutiérrez-Estrada, "Use of artificial neural networks in near-infrared reflectance spectroscopy calibrations for predicting the inclusion percentages of wheat and sunflower meal in compound feedingstuffs," *Appl. Spectrosc.*, vol. 60, no. 9, pp. 1062–1069, Sep. 2006.
[40] J. Workman Jr. and L. Weyer, *Practical Guide to Interpretive Near-Infrared Spectroscopy*. Boca Raton, FL, USA: CRC Press, 2007.
[41] B. G. Osborne, T. Fearn, and P. H. Hindle, *Practical NIR Spectroscopy: With Applications in Food and Beverage Analysis*, 2nd ed. Singapore: Longman, 1993, p. 227.
[42] T. Rutpralom, P. Kumhom, and K. Chamnongthai, "Nondestructive maturity determination of durian by using microwave moisture sensing," in *Proc. IEEE Int. Conf. Ind. Technol. (ICIT)*, Bangkok, Thailand, Dec. 2002, pp. 155–158.
[43] R. B. Jordan, E. F. Walton, K. U. Klages, and R. J. Seelye, "Postharvest fruit density as an indicator of dry matter and ripened soluble solids of kiwifruit," *Postharvest Biol. Technol.*, vol. 20, no. 2, pp. 163–173, Sep. 2000.
[44] S. Hor, M. Léchaudel, H. Mith, and C. Bugaud, "Fruit density: A reliable indicator of sensory quality for mango," *Scientia Horticulturae*, vol. 272, Oct. 2020, Art. no. 109548.

AMORNRIT PUTTIPIPATKAJORN received the B.S. degree in electrical engineering from Kasetsart University, Bangkok, in 1997, and the M.S. and Ph.D. degrees in electrical engineering from the University of Montpellier II, Montpellier, France, in 2000 and 2005, respectively. He is currently an Associate Professor with Kasetsart University. His research interests include machine learning, artificial intelligence, non-destructive techniques, and spectroscopy applications in agriculture.

ANUPUN TERDWONGWORAKUL received the B.E. degree in agricultural engineering from Kasetsart University, Nakhon Pathom, Thailand, in 1986, and the M.S. degree in welding technology and the Ph.D. degree in agricultural engineering from Cranfield University, Cranfield, U.K., in 1990 and 1995, respectively.

He is currently a Professor with the Faculty of Engineering at Kamphaeng Saen, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom. His current research interest includes non-destructive evaluation of internal quality of agricultural produce.

AMORNDEJ PUTTIPIPATKAJORN received the B.S. degree in food engineering from Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom, Thailand, in 1997, and the M.S. degree in chemical engineering from Kasetsart University, Bangkok, Thailand, in 2001. He is currently an Associate Professor with Kasetsart University. His research interests include the spectroscopy and spectral imaging analysis, and its applications in food and agriculture.

SUPACHAI KULMUTIWAT received the B.S. and M.S. degrees in mechanical engineering from Kasetsart University, Bangkok, in 1998 and 2004, respectively. He is currently a Lecturer with the Faculty of Engineering at Kamphaeng Saen, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom, Thailand. His current research interests include autonomous agricultural machinery and precision farming.

PEERAPONG SANGWANANGKUL received the B.S. (Hons.) and M.S. degrees in horticulture from Kasetsart University, Bangkok, Thailand, in 1994 and 1998, respectively, and the Ph.D. degree in horticulture from the University of Hawai'i at Manoa, USA, in 2007.

He is currently a Professional Researcher with the Postharvest Technology Center, Faculty of Agriculture at Kamphaeng Saen, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom, Thailand. His research interests include preharvest factors affecting quality of horticultural commodities and postharvest handling and technology for tropical fruits and vegetables.

THANA CHEEPSOMSONG received the B.Eng. and M.Eng. degrees in mechanical engineering from Kasetsart University, Thailand, in 2000 and 2003, respectively, and the Ph.D. degree in engineering from the University of Sussex, U.K., in 2014.

Since 2003, he has been a Lecturer with the Faculty of Engineering at Kamphaeng Saen, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom, Thailand. His current research interests include 3D vision, robotics, and applications of AI technology.