EASR

Engineering and Applied Science Research

https://www.tci-thaijo.org/index.php/easr/index

Published by the Faculty of Engineering, Khon Kaen University, Thailand

Enhancing of the rapid evaluation of sugarcane energy content for energy cane varieties selection purposes in breeding program using Near-Infrared spectroscopy (NIRS)

Kantisa Phoomwarin[1)], Jetsada Posom[1, 2)], Khwantri Saengprachatanarug[1, 3)] and Arthit Phuphaphud*[1, 3)]

[1)]Faculty of Engineering, Khon Kaen University, Khon Kaen 40002, Thailand
[2)]Center for Alternative Energy Research and development, Khon Kaen University, Khon Kaen 40002, Thailand
[3)]Northeast Thailand Cane and Sugar Research Center [NECS], Khon Kaen University, Khon Kaen 40002, Thailand

Received 28 August 2023
Revised 25 November 2023
Accepted 28 November 2023

Abstract

The integration of Near-Infrared spectroscopy (NIRs) as a predictive tool for rapidly assessing the gross calorific value (GCV) of dried shredded sugarcane addresses a significant bottleneck in sugarcane breeding programs. A total of 110 samples were collected, and a lab-type Fourier Transform near infrared (FT-NIR) spectrometer operating across a wavenumber range of 12,800 to 3,600 cm^{-1} was utilized for spectral acquisition. This study explored various variable selection algorithms with partial least squares (PLS) regression for model development. Variable selection methods included variable importance in projection (VIP), successive projections algorithm (SPA), genetic method (GA), and correlation (r). The r-PLS algorithm along with standard normal variate and first derivative (SNV + D^1) pre-treatment yielded a highly effective model exhibiting a coefficient of determination in prediction (R^2_P) of 0.88 and the root mean square error of prediction ($RMSE_P$) of 278.0 J/g. This study showed that the NIRs could provide a reliable method for rapid, fairly accurate, and precise GCV estimation. However, future research should investigate the use of additional modeling algorithms, such as non-linear regression, to improve model accuracy and utilize sample data with an evenly spread range of GCV values to validate the model performance with greater confidence.

Keywords: Sugar mill, Heating values, Fourier transform near infrared spectroscopy, Energy characteristics, Rapid measurement

1. Introduction

Sugarcane is increasingly being utilized in various applications. For instance, many sugar factories use it to produce sugar as the main product and utilize bagasse, the residual material left after juice extraction, to generate heat and electricity [1]. This method is widely adopted and has achieved great success. On the other hand, some power plants are using sugarcane to produce electricity by directly burning dried chopped cane stalks as a new sustainable energy source from biomass plants. This approach is a promising alternative energy source as fossil fuels may not last beyond mid-century and lead to an imbalance of greenhouse gases [2]. Therefore, it is crucial to develop improved sugarcane varieties through breeding programs that are suitable for specific purposes.

Roach [3] outlined three stages in sugarcane improvement breeding programs. The first stage concentrates on the selection and crossbreeding of primary variety clones. The second stage focuses on creating interspecific hybrids, and the third stage involves exploiting these interspecific hybrids. These three phases often take over ten years to complete.

During each stage, the evaluation of sugarcane energy content of each clone is important because this parameter plays an important role in clone selection processes. The energy content can be estimated using the higher heating value (HHV) or gross calorific value (GCV) [4]. The conventional method used for determining the GCV is bomb calorimetry. This method is costly and time-consuming and can be determined only in a small amount of sample for each analysis. As a result, breeders can only select a limited number of clones to measure the GCV and discard some clones that may have good energy content.

The application of Near-Infrared spectroscopy (NIRs) has been widely used for determining the energy content of agricultural products [4-8]. This technology provides a rapid assessment, which can be performed without any sample preparation. In the context of breeding programs, NIR spectroscopy has been successfully employed by Baillères et al. [9] and Yeh et al. [10] to quickly evaluate major wood characteristics in eucalyptus and pine, respectively. These models have proven effective in enabling the efficient screening of large breeding populations for a wide range of phenotypic traits. Furthermore, the wavenumbers from the Near-infrared (NIR) region have been utilized in drone imagery for the detection of sugarcane canopy [11].

Consequently, this research aimed to explore the potential of utilizing the NIRs to evaluate the GCV of sugarcane as a tool for rapidly screening a large breeding population for any new species. Additionally, a comparison of model accuracy was carried out between using the entire wavelength range and the relevant wavelengths selected via the VIP, SPA, GA, and r methods. If this technology can provide a good result, it will be a helpful tool in breeding programs, resulting in breeders can choose more numbers of clones for energy content assessment and increase the chances of obtaining sugarcane clones with good energy characteristics.

*Corresponding author.
Email address: arthphu@kku.ac.th
doi: 10.14456/easr.2024.7

2. Materials and methods

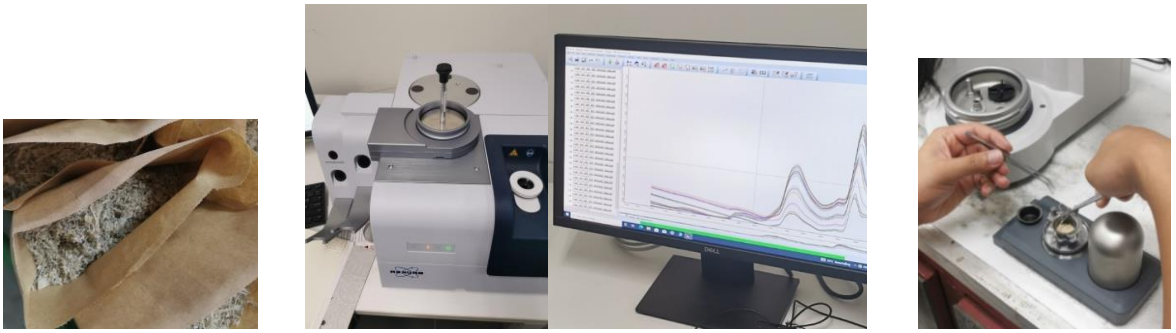
2.1 Sample preparation

In this study, sugarcane samples were provided by the Khon Kaen Field Crop Research Centre in Thailand. These samples were from an energy-focused sugarcane breeding program. To ensure the robustness of the model for use in breeding programs, six sugarcane varieties were included: F03-362, KK09-0358, KK09-0939, TPJ04-768, KK3, and UT12. The KK3 and UT12 varieties which were the primary commercial cultivars in Thailand and serve as essential references for comparison to any new cultivars. The remaining varieties were selected for their high potential as energy cane.

The experiment was conducted every two months once the sugarcane has grown to 8-14 months. In order to establish a comprehensive model, a range of energy characteristic values were obtained by collecting samples from different varieties and at various stages of maturity. The following procedure were used for preparing the samples:

1. Separate the sugarcane leaves from the stalk.
2. Divide each stalk into three sub-sections: top, middle, and bottom sections.
3. Grind each sub-section sample into shredded sample.
4. Extract juice from the shredded sample.
5. Boil and dehydrate the sample from previous step.
6. Dry all samples in an oven (UN55, Memmert, Germany) at 105 ºC until they reach a constant mass.

After completing the aforementioned steps, the dried shredded samples (Figure 1a) were subjected to spectral scanning, carried out at an ambient temperature of approximately 25 ˚C (Figure 1b). Subsequently, the GCV of all samples was determined using a bomb calorimeter (C5003, IKA, Germany), as illustrated in Figure 1c.



a **b** **c**

Figure 1 (a) Dried shredded sample, (b) Spectral acquisition using Fourier transform infrared (FT-NIR) spectrometer, and (c) GCV determination using a bomb calorimeter

2.2 NIR scanning

Lab-type spectrometer, FT-NIR spectrometer (MPA II, Bruker, Germany) across a wavenumber range of 12,800 to 3,600 cm^{-1}, was used for spectral acquisition of dried shredded sample in a quartz bottom sampling cup (87 mm in diameter and 87.5 mm in height) and repacked by hand between each scan. The FT-NIR spectrometer was used in the diffuse reflectance mode at room temperature (25 ± 2 ˚C). Before each sample was scanned, a gold plate was scanned as a reference. A sample cup was placed on a rotary sample holder and scanned through the quartz at the bottom of the cup. The sample was scanned 64 times with a resolution of 16 cm^{-1}, and were averaged.

2.3 Reference method

After the scanning process, the scanned samples were milled into small particles by universal milling machine (M 20 Universal mill, IKA, Germany) to determine the GCV. The GCV determination was conducted using an automated isoperibol calorimeter (C5003, IKA, Germany). Approximately 0.4 g of the milled sample was placed into a combustion polyethylene bag (C12, IKA®-Werke, gross calorific value of 46463 J/g, size 4 × 3.5 mm) and combusted within the calorimeter. Prior to the measurements, pelletized benzoic acid was used to calibrate the bomb calorimeter. Furthermore, a cubic centimeter of water was introduced into the calorimeter vessel to endorse that the bomb's interior was saturated with water vapor at both the start and the conclusion of the experiment [12]. This GCV determination was carried out under ambient conditions with a room temperature of approximately 25 °C. The water in the outer vessel, commonly known as the calorimeter jacket, was maintained at 25 °C. This water jacket helped ensure that the system reached thermal equilibrium quickly after the sample was ignited and helped minimize heat loss to the surroundings.

Each sample underwent duplicate GCV determinations. The precision of the reference values was assessed by calculating the repeatability value (Rep), which can be computed from the standard deviation of the difference values between duplicates and the standard error of the laboratory (SEL) using the following Equation [13]:

$$\text{SEL} = \sqrt{\frac{\sum_i(\sum_j(y_{ij}-\bar{y}_i)^2)}{N}} \quad (1)$$

Here, the terms "$\bar{y}_i$ " denote the reference method mean value of all replicates of the *i*th sample, "y_{ij}" is the reference value of the *j*th replicate of the *i*th sample, while "*N*" signifies the total number of samples, and *i* is the number of a sample.

2.4 Spectral pre-processing and model development

Figure 2 illustrates the comprehensive experimental setup of this research. After collecting the reference value (dependent variable), i.e., GCV, and corresponding spectra (independent variable), the relationship between them was performed. Prior to modeling, the NIR spectra were pretreated using the following techniques, first derivative (D^1), second derivative (D^2), standard normal variate (SNV), baseline offset, mean centering, and SNV + D^1. The following pre-processing methods were implemented to mitigate the potential impacts of both multiplicative scattering and baseline shifts, which could emerge from phenomena like light scattering or variations in the lengths of light paths within the samples, as observed in previous research [5, 14]. Additionally, wavelength selection methods, such as variable importance in projection (VIP), successive projections algorithm (SPA), genetic method (GA), and correlation (r), were employed to reduce the number of independent variables and identify the significant independent variables.

VIP is a variable selection algorithm that can be used to identify the most important variables in a multivariate data set. It is based on partial least squares (PLS) regression, which is a statistical method for modeling relationships between multiple predictor variables and a response variable. SPA is based on the idea of projecting the data onto a series of orthogonal planes, and then selecting the variables that are most important in each plane. GA is a search heuristic inspired by the process of natural selection. It works by evolving a population of solutions to a problem over multiple generations, selecting the best solutions from each generation and combining them to create new solutions. Maraphum et al. [15] said that the GA was a flexible technique for variable selection in multivariate calibration. Lastly, r is a fundamental statistical concept used to measure the linear relationship between two variables. The most used correlation algorithm is the Pearson correlation coefficient, which was applied to this research.

In this study, the entire sample set was subject to random division into two distinct subsets by taking one of every five samples, with the reference data sorted in ascending order: a calibration set and a prediction set, both serving modeling purposes. The calibration set consisted of 80% of the total samples, with the remaining 20% allocated to the prediction set. To ensure the creation of robust models, the minimum and maximum of reference values were included in the calibration set. For modeling, the PLS regression was utilized with leave-one-out cross-validation to prevent overfitting [16]. The pre-processing and modeling methods were executed using a custom in-house MATLAB code (The MathWorks, Inc., Natick, Massachusetts, USA, license no: 40846673).

2.5 Model performance evaluation

The model's performance was evaluated using the following indices: coefficient of determination in calibration (R^2), coefficients of determination in cross-validation (R^2_{cv}), coefficients of determination in prediction (R^2_p), root mean square error of calibration ($RMSE_C$), root mean square error of cross-validation ($RMSE_{CV}$), root mean square error of prediction ($RMSE_P$), Bias, and the ratio of prediction to deviation (RPD). To determine the appropriate PLS model for each case, including spectral pre-processing techniques, number of PLS factors (F), the wavelength selection method, and regression coefficients. These were achieved by seeking the lowest $RMSE_{CV}$ value [17].

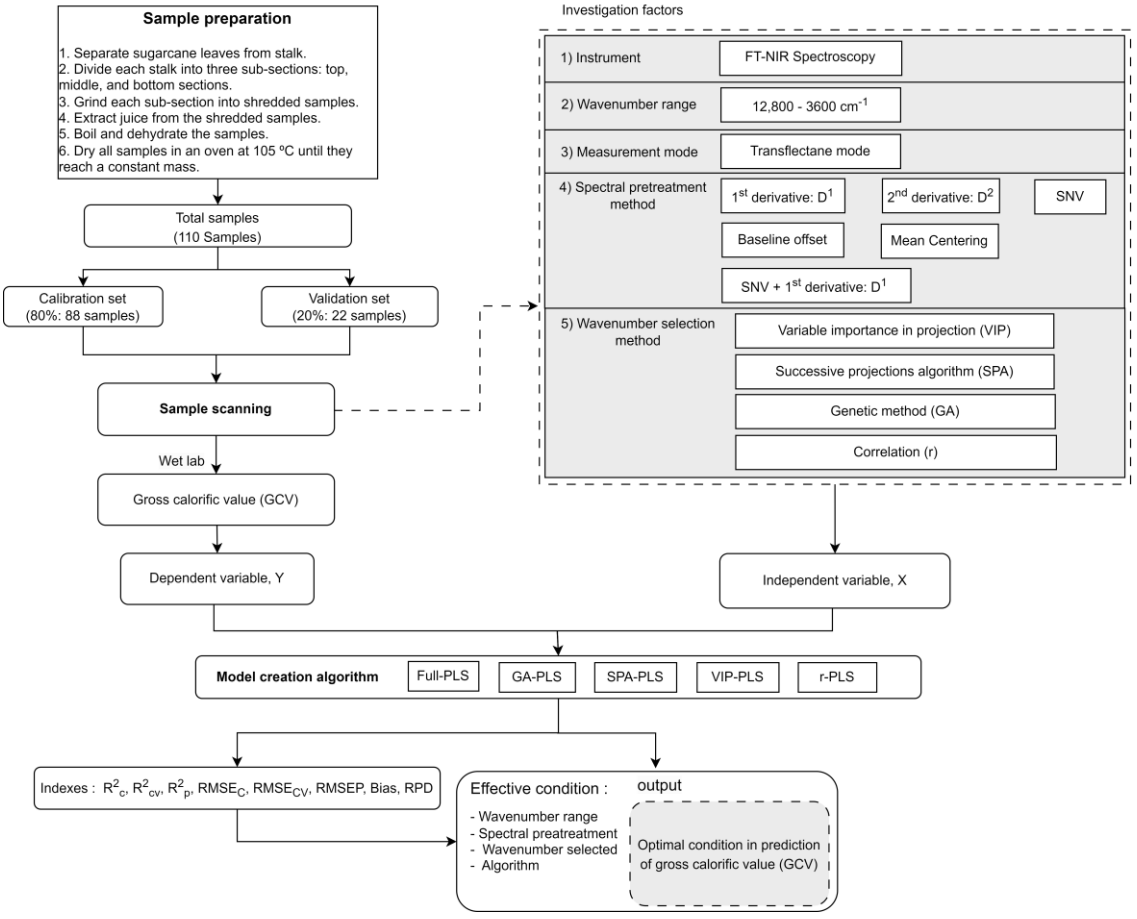


Figure 2 Experimental overview, including sample preparation, investigation factor, spectral treatment method, model creation algorithm, and output

3. Results and discussion

3.1 NIR Spectra and statistical characteristics of the GCV of dried shredded cane

Figure 3a displays the raw absorbance spectra of dried shredded cane stalks, exhibiting a comparable overall structure across all samples. Each color line indicates spectra obtained from each sugarcane variety. The spectral absorbance profiles of the six sugarcane varieties were comparable, with no significant differences between any varieties. However, there was a noticeable baseline shift that appeared to separate the spectra into two distinct groups. The observed spectra could be categorized into two distinct groups, indicating that the samples in each group underwent different treatments in breeding programs. Specifically, the group with lower spectral absorption was associated with drought-tolerant sugarcane, whereas the other group was linked to irrigated sugarcane. Figure 3b presents the spectra subjected to pre-treatment using the SNV + D^1 technique, which was found to yield the optimal model in this study. The pre-treated spectra demonstrated the ability to eliminate spectral shifts that occur in raw spectra.

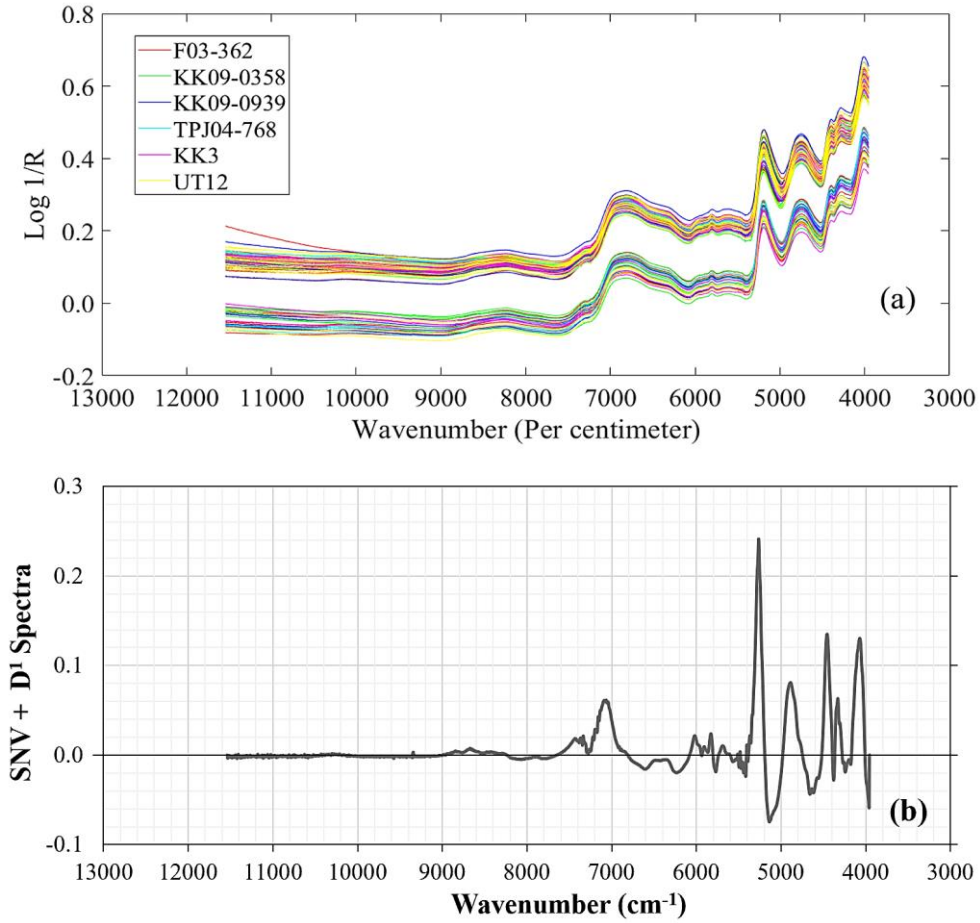


Figure 3 Spectra of 110 dried shredded cane samples: (a) raw spectra; (b) SNV + first derivative spectra

Table 1 displays the statistical characteristics of the reference data used for model development, including minimum, maximum, mean, and standard deviation. The range of GCV values falls between 5,977.0 and 17,905.0 J/g, which was wider compared to previous research by Phuphaphud et al. [8], where the heating values of fibrous cane ranged between 15,379.3 and 17,545.9 J/g, and another study by Posom et al. [7] with GCV of bagasse ranging from 16,278.4 to 17,802.1 J/g. The broader range of GCV in this experiment resulted from the inclusion of sugarcane samples at various stages of maturity and from different cultivars.

Based on the precision assessment of the reference method, the Rep value was calculated to be 246.9 J/g, while the SEL value was determined to be 119.3 J/g. These outcomes confirm the accuracy of the reference data, as the comparatively low SEL value with the standard deviation values indicates an acceptable level of precision.

Table 1 Statistical characteristics and precision values of laboratory for gross calorific value (GCV, Unit: J/g) samples

Data sets	Number of samples	Maximum	Mean	Minimum	Standard deviation	Rep	SEL
Calibration set	88	17,905.0	15,126.8	5,977.0	3,665.0	246.9	119.3
Prediction set	22	17,899.0	14,958.5	6,043.0	3,857.4		

Rep: repeatability of reference data.
SEL: standard error of laboratory.

3.2 PLS models of GCV prediction

Table 2 displays the results of GCV prediction, consisting of the algorithm, selected wavenumber, pre-processing techniques, F, R^2_C, $RMSE_C$, R^2_{cv}, $RMSE_{CV}$, r, $RMSE_P$, Bias, and RPD. This article involved comparing the model's performance constructed using different PLS algorithms and pre-treatment techniques. The PLS algorithms evaluated included Full-PLS, GA-PLS, SPA-PLS, VIP-PLS, and r-PLS methods. Each algorithm utilized spectra from seven pre-processing methods: raw spectra, D^1, D^2, SNV, baseline offset, mean centering, and SNV + D^1. Based on the results, the utilization of r-PLS with SNV+D^1 provided the best model for GCV prediction, employing eight latent variables. The total wavenumbers used in the optimal model were reduced from 2662 to 1300 wavenumbers. The R^2_c, R^2_p, $RMSE_C$, $RMSE_P$, Bias, and RPD values for the optimal GCV prediction model were 0.98, 0.88, 490.8 J/g, 1285.3 J/g, 278.0 J/g, and 3.00, respectively. The R^2_p value of the optimal model in this research closely resembled the study conducted by Sirisomboon et al. [5], where they measured the GCV of bamboo chips and obtained an R^2_p, RMSEP, Bias, and RPD of 0.84, 150 J/g, 1.81 J/g, and 2.52. It also showed similar results to another study by Posom et al. [18], where they measured the GCV of ground bamboo and obtained an R^2_p, RMSEP, Bias, and RPD of 0.92, 122 J/g, 14.4 J/g, and 3.66.

Williams et al. [13] recommended that models with an R^2 of 0.83–0.90 are usable with caution for most applications, including research. Additionally, Zornoza et al. [19] advised that predictions with an R^2 >0.81 are reasonable, and predictions with an RPD between 2.5 and 3.0 can be classified as good. Based on these guidelines, the models constructed in this study can be confidently utilized for the rapid estimation of GCV in dried shredded cane. As a result, these findings suggest that NIR spectroscopy can serve as a valuable tool in breeding programs, aiding breeders in the assessment of sugarcane's energy content. Consequently, breeders could potentially assess a higher number of clones for energy content and thereby enhance the likelihood of identifying sugarcane clones with favorable energy characteristics.

Table 2 Prediction results of gross calorific value (GCV)

Algorithm	Selected wavelen-umber	Pre-processing	Calibration set					Prediction set			
			F	R^2_c	$RMSE_C$ (J/g)	R^2_{cv}	$RMSE_{CV}$ (J/g)	R^2_p	$RMSE_P$ (J/g)	Bias (J/g)	RPD
Full-PLS	2662	Raw spectra	14	0.91	1077.1	0.77	1754.2	0.68	2171.5	-94.9	1.78
		D^1	8	0.98	484.4	0.87	1318.0	0.86	1376.7	422.4	2.80
		D^2	6	0.97	612.6	0.87	1328.9	0.84	1502.2	338.4	2.57
		SNV	12	0.85	1434.0	0.68	2099.3	0.64	2297.6	-108.2	1.68
		Baseline offset	14	0.90	1156.2	0.77	1754.4	0.61	2393.2	-116.7	1.61
		Mean Centering	15	0.93	988.5	0.79	1689.5	0.63	2352.7	-52.1	1.64
		SNV + D^1	9	0.99	368.3	0.86	1367.4	0.88	1308.9	252.0	2.95
GA-PLS	1640	Raw spectra	14	0.92	1056.2	0.79	1688.4	0.69	2158.5	-71.5	1.79
	1580	D^1	8	0.98	514.8	0.87	1317.2	0.85	1442.5	388.7	2.67
	740	D^2	6	0.97	589.9	0.87	1311.4	0.79	1692.1	503.1	2.28
	980	SNV	12	0.85	1402.1	0.60	2400.7	0.61	2397.8	-191.7	1.61
	1740	Baseline offset	13	0.88	1253.0	0.74	1841.7	0.58	2484.4	-188.4	1.55
	920	Mean Centering	14	0.91	1101.2	0.77	1739.7	0.63	2327.0	-224.7	1.66
	500	SNV + D^1	8	0.97	675.1	0.69	2038.0	0.83	1569.7	64.7	2.46
SPA-PLS	420	Raw spectra	15	0.85	1409.0	0.68	2101.0	0.73	1967.8	444.1	1.96
	540	D^1	9	0.91	1084.6	0.84	1472.1	0.74	1940.2	188.2	1.99
	2000	D^2	6	0.97	620.4	0.87	1335.3	0.84	1502.5	333.1	2.57
	1500	SNV	15	0.92	1009.7	0.79	1680.9	0.69	2133.0	0.9	1.81
	1960	Baseline offset	15	0.87	1324.4	0.74	1852.5	0.53	2612.9	-322.0	1.48
	1480	Mean Centering	15	0.92	1035.0	0.79	1682.0	0.69	2153.4	65.9	1.79
	400	SNV + D^1	9	0.90	1144.1	0.87	1312.7	0.76	1905.1	75.5	2.02
VIP-PLS	943	Raw spectra	10	0.83	1493.3	0.60	2328.6	0.67	2203.4	346.2	1.75
	670	D^1	6	0.94	880.5	0.87	1301.3	0.82	1616.4	43.9	2.39
	566	D^2	6	0.96	752.4	0.89	1211.2	0.81	1646.7	287.4	2.34
	991	SNV	14	0.92	1062.3	0.83	1506.4	0.65	2278.2	-141.9	1.69
	1022	Baseline offset	12	0.85	1417.3	0.76	1784.3	0.63	2360.3	-43.6	1.63
	848	Mean Centering	15	0.92	1005.5	0.82	1559.6	0.68	2186.9	92.0	1.76
	649	SNV + D^1	8	0.98	575.6	0.88	1268.7	0.83	1585.1	171.4	2.43
r-PLS	360	Raw spectra	11	0.64	2190.4	0.51	2606.2	0.60	2432.0	25.0	1.59
	1220	D^1	8	0.98	495.5	0.89	1230.9	0.84	1488.7	339.6	2.59
	880	D^2	7	0.98	528.4	0.91	1148.5	0.81	1661.1	118.7	2.32
	560	SNV	15	0.92	1060.3	0.80	1625.4	0.71	2038.2	-358.3	1.89
	1940	Baseline offset	13	0.81	1586.7	0.76	2267.2	0.53	2626.0	-175.5	1.47
	1940	Mean Centering	15	0.95	824.5	0.83	1493.4	0.71	2049.8	220.0	1.88
	1300	**SNV + D^1**	**8**	**0.98**	**490.8**	**0.88**	**1337.3**	**0.88**	**1285.3**	**278.0**	**3.00**

F: number of PLS factors.
R^2_c: coefficient of determination in calibration.
R^2_{cv}: coefficient of determination in cross-validation.
R^2_p: coefficient of determination in prediction.
$RMSE_C$: the root mean square error of calibration.
$RMSE_{CV}$: the root mean square error of cross-validation.
$RMSE_P$: the root mean square error of prediction.
Bias: the average of difference values between observed values and predicted values.
RPD: the ratio of prediction to deviation.

In Figure 4, scatter plots illustrate the correlation between the actual values and the predicted values of the GCV of shredded cane. These predicted values were generated using the optimized GCV prediction model, established by employing spectra pre-processed through a combination of the SNV and D^1 methods, and then modeled using the r-PLS algorithm. Although using sugarcane samples from various stages of maturity and different cultivars provided a wide range of GCV values, the GCV was not evenly distributed and was instead concentrated into two distinct groups. This uneven distribution may have made it difficult to assess the model performance, especially for the R^2_p value. Additionally, using the models developed in this study may lead to low accuracy in predictions for samples with GCV values in the range of 9,000 to 14,000 J/g, as shown in Figure 4a. Therefore, future studies should use sample data with a more evenly distributed range of GCV values to avoid low accuracy in model predictions for some samples and to confirm the model performance with greater confidence.

A preview of the average SNV + D^1 spectra is presented in Figure 5. The red points highlight wavenumbers selected through the r method for modeling purposes, based on their correlation with GCV and spectral data. Figure 6 illustrates the regression coefficient plots of the effective GCV prediction model. This model was developed using the SNV + D^1 spectra in combination with the r-PLS algorithm. A pronounced regression coefficient at a particular wavenumber suggests that the vibration of the corresponding bond at that frequency could have a substantial influence on the model's prediction [4]. Highlighting wavenumbers with significant regression coefficients, the peaks situated around 10941 (914 nm), 10183 (982 nm), 4288 (2332 nm), and 4032 (2480 nm) cm^{-1} correspond to dry cellulose [13]. Peaks 9259 (1080 nm), 7168 (1395 nm), and 5208 (1920 nm) cm^{-1} are related to the structure of benzene, CH_2, and CONH, respectively [20]. Peaks were also evident at 10661 (938 nm) cm^{-1}, attributed to the CH_2 structure (C–H stretching the third overtone) [21].

CH_2 is the chemical formula for the methylene group, which is found in many organic molecules, including cellulose [22]. Cellulose is composed of glucose units linked together in long chains. When combusted, these glucose units release energy, contributing to the GCV of the material. Lin et al. [23] described cellulose as a flexible material for energy storage and conversion, indicating that more cellulose in sugarcane can lead to more energy content or GCV. The CONH is the chemical formula for the amide group [24], found in many organic molecules, including proteins and carbohydrates. Although proteins and carbohydrates may not have a significant impact on the GCV of a material, changes in their composition can lead to variation in GCV. Therefore, the wavenumbers related to these compositions played an important role in model prediction.

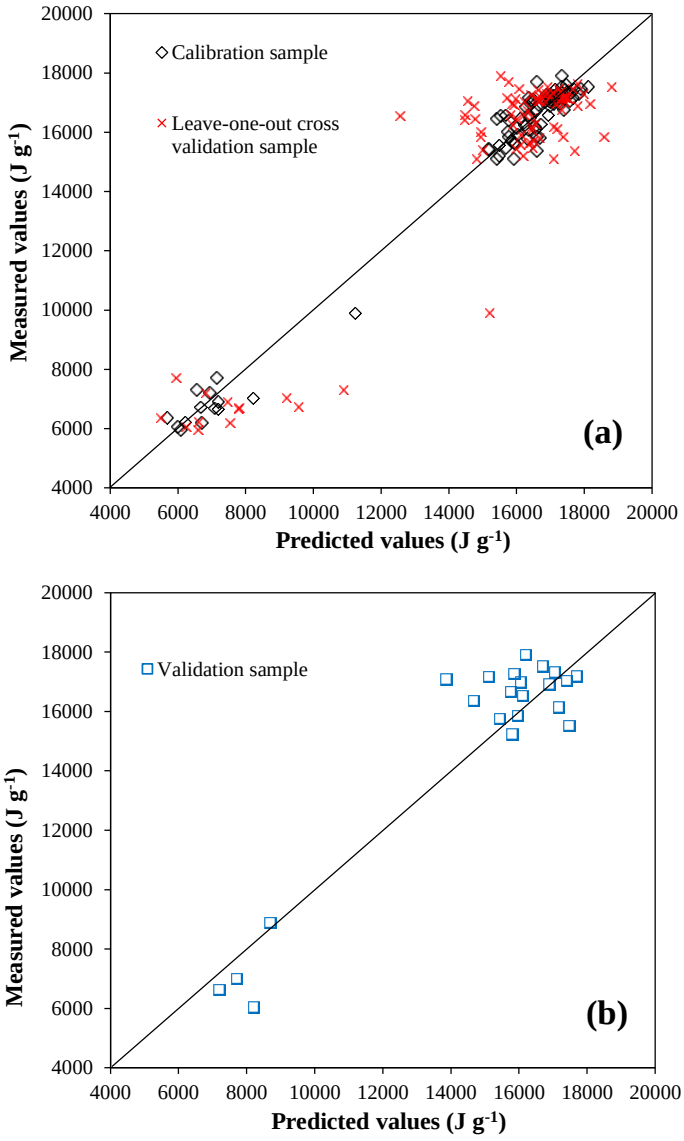


Figure 4 NIR spectroscopy predicted gross calorific value (GCV) using r-PLS with SNV+D^1 versus actual values from the reference method in the calibration (a) and prediction (b) models.

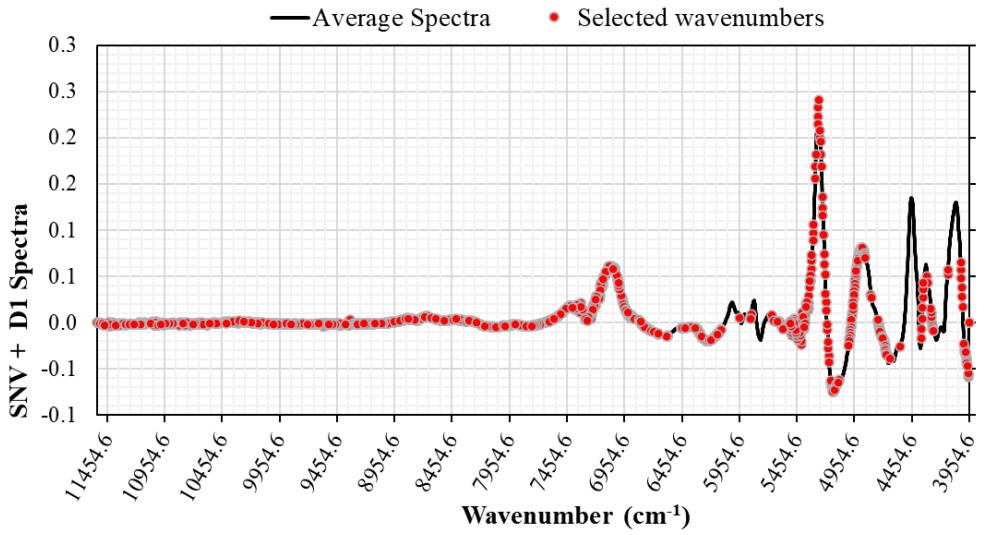


Figure 5 Wavelength spectra and interval wavelength selected by correlation (r) algorithm

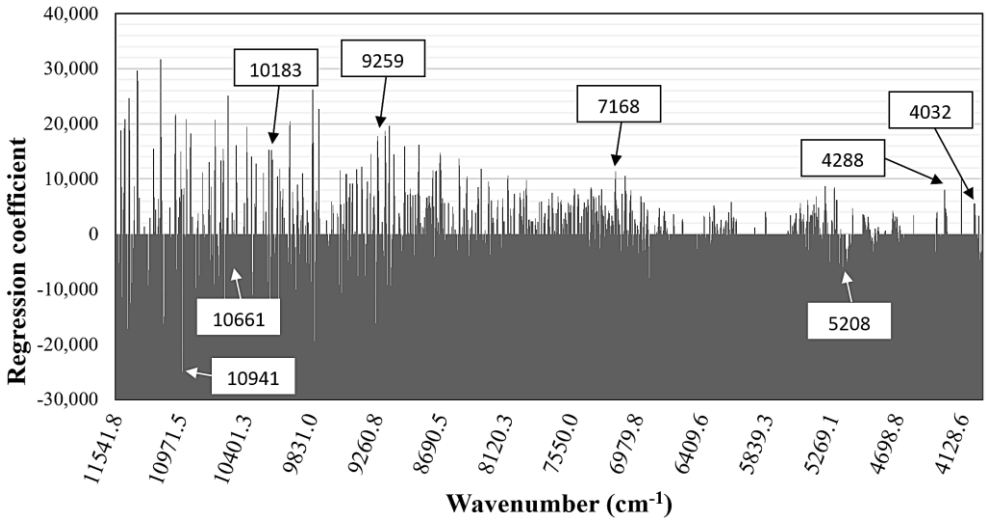


Figure 6 Regression coefficient for GCV model developed using r-PLS regression coupled with spectral pre-treatment of SNV+D[1]

4. Conclusions

The integration of Near-Infrared spectroscopy (NIRs) as a tool for predicting the gross calorific value (GCV) of dried shredded cane presents fairly reliable method for rapidly evaluating sugarcane energy content but the better modeling algorithm have to be explored for example non-linear modeling algorithm, addressing a critical bottleneck in breeding programs. The research explored various algorithms, pre-processing techniques, and spectral analyses to identify the most effective model for GCV prediction. The adoption of the r-PLS algorithm in conjunction with SNV + D[1] pre-treatment emerged as the optimal choice, providing accurate predictions with fewer latent variables and selected wavenumbers. Furthermore, the analysis of significant wavenumbers and regression coefficients elucidated the underlying molecular structures associated with GCV, contributing to a deeper understanding of the relationship between spectral data and energy content. This knowledge can guide future research and enhance the interpretation of spectral data in various applications.

5. Acknowledgements

This research is supported by Research and Graduate Studies, Khon Kean University. Additionally, the authors extend their gratitude to the Northeast Thailand Cane and Sugar Research Centre, Faculty of Agriculture, Khon Kaen University, for granting access to their instruments. Special appreciation is directed towards the Khon Kaen Field Crop Research Centre of Khon Kaen, Thailand, for their provision of valuable sugarcane samples.

6. References

[1] Bocci E, Di Carlo A, Marcelo D. Power plant perspectives for sugarcane mills. Energy. 2009;34(5):689-98.
[2] Carvalho-Netto OV, Bressiani JA, Soriano HL, Fiori CS, Santos JM, Barbosa GV, et al. The potential of the energy cane as the main biomass crop for the cellulosic industry. Chem Biol Technol Agric.2014;1:1-8.
[3] Roach BT. Origin and improvement of the genetic base of sugarcane. Proceeding of Australian Society of Sugar Cane Technologists; 1989 May 2-5; Tweed Heads, Australia. p. 34-47.

[4] Nakawajana N, Posom J, Paeoui J. Prediction of higher heating value, lower heating value and ash content of rice husk using FT-NIR spectroscopy. Eng J. 2018;22(5):45-56.
[5] Sirisomboon P, Funke A, Posom J. Improvement of proximate data and calorific value assessment of bamboo through near infrared wood chips acquisition. Renew Energy. 2020;147:1921-31.
[6] Posom J, Shrestha A, Saechua W, Sirisomboon P. Rapid non-destructive evaluation of moisture content and higher heating value of *Leucaena leucocephala* pellets using near infrared spectroscopy. Energy. 2016;107:464-72.
[7] Posom J, Phuphaphud A, Saengprachatanarug K, Maraphum K, Saijan S, Pongkan K, et al. Real-time measuring energy characteristics of cane bagasse using NIR spectroscopy. Sens Bio-Sens Res. 2022;38:100519.
[8] Phuphaphud A, Saengprachatanarug K, Posom J, Taira E, Panduangnate L. Prediction and classification of energy content in growing cane stalks for breeding programmes using visible and shortwave near infrared. Sugar Tech. 2022;24(5):1497-509.
[9] Baillères H, Davrieux F, Ham-Pichavant F. Near infrared analysis as a tool for rapid screening of some major wood characteristics in a eucalyptus breeding program. Ann For Sci. 2002;59:479-90.
[10] Yeh TF, Yamada T, Capanema E, Chang HM, Chiang V, Kadla JF. Rapid screening of wood chemical component variations using transmittance near-infrared spectroscopy. J Agric Food Chem. 2005;53(9):3328-32.
[11] Chea C, Saengprachatanarug K, Posom J, Wongphati M, Taira E. Sugarcane canopy detection using high spatial resolution UAS images and digital surface model. Eng Appl Sci Res. 2019;46(4):312-7.
[12] Jessup RS. Precise measurement of heat of combustion with a bomb calorimeter. Washington: National Bureau of standard; 1960.
[13] Williams P, Antoniszyn J, Manley M. Near infrared technology: getting the best out of light. Stellenbosc: African Sun Media; 2019.
[14] Taira E, Ueno M, Furukawa N, Tasaki A, Komaki Y, Nagai JI, et al. Networking system employing near infrared spectroscopy for sugarcane payment in Japan. J Near Infrared Spectrosc. 2013;21(6):477-83.
[15] Maraphum K, Ounkeaw A, Kasemsiri P, Hiziroglu S, Posom J. Wavelengths selection based on genetic algorithm (Ga) and successive projections algorithms (spa) combine with pls regression for determination the soluble solids content in nam-dokmai mangoes based on near infrared spectroscopy. Eng Appl Sci Res. 2022;49(1):119-26.
[16] Sammut C, Webb GI, editors. Leave-One-Out Cross-Validation. Encyclopedia of Machine Learning. Boston: Springer; 2010. p. 600-1.
[17] Sirisomboon P, Posom J. On-line measurement of activation energy of ground bamboo using near infrared spectroscopy. Renew Energy. 2019;133:480-8.
[18] Posom J, Sirisomboon P. Evaluation of lower heating value and elemental composition of bamboo using near infrared spectroscopy. Energy. 2017;121:147-58.
[19] Zornoza R, Guerrero C, Mataix-Solera J, Scow KM, Arcenegui V, Mataix-Beneyto J. Near infrared spectroscopy for determination of various physical, chemical and biochemical properties in Mediterranean soils. Soil Biol Biochem. 2008;40(7):1923-30.
[20] Workman J, Weyer L. Practical guide to interpretive near-infrared spectroscopy. Boca Raton: CRC Press; 2007.
[21] Wang J, Wang J, Chen Z, Han D. Development of multi-cultivar models for predicting the soluble solid. Postharvest Biol Technol. 2017;129:143-51.
[22] Filho GR, de Assunção RMN, Vieira JG, da S Meireles C, Cerqueira DA, da Silva Barud H, et al. Characterization of methylcellulose produced from sugar cane bagasse cellulose: crystallinity and thermal properties. Polym Degrad Stab. 2007;92(2):205-10.
[23] Lin Z, Li S, Huang J. Natural cellulose substance based energy materials. Chem Asian J. 2021;16(5):378-96.
[24] Hu L, Xu S, Zhao Z, Yang Y, Peng Z, Yang M, et al. Ynamides as racemization-free coupling reagents for amide and peptide synthesis. J Am Chem Soc. 2016;138(40):13135-8.