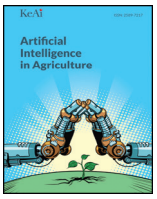
Artificial Intelligence in Agriculture 6 (2022) 266–275

Contents lists available at ScienceDirect

Artificial Intelligence in Agriculture

journal homepage: http://www.keaipublishing.com/en/journals/artificial-intelligence-in-agriculture/



Predicting the true density of commercial biomass pellets using near-infrared hyperspectral imaging

Lakkana Pitak [a], Khwantri Saengprachatanarug [a], Kittipong Laloon [a], Jetsada Posom [a,b,*]

[a] Department of Agricultural Engineering, Faculty of Engineering, Khon Kaen University, Khon Kaen 40002, Thailand
[b] Center for Alternative Energy Research and Development, Khon Kaen University, Thailand

ARTICLE INFO

Article history:
Received 25 July 2022
Received in revised form 14 November 2022
Accepted 15 November 2022
Available online 17 November 2022

Keywords:
True density
Hyperspectral imaging
Biomass pellet
Variable selection method

ABSTRACT

The use of biomass is increasing because it is a form of renewable energy that provides high heating value. Rapid measurements could be used to check the quality of biomass pellets during production. This research aims to apply a near-infrared (NIR) hyperspectral imaging system for the evaluation of the true density of individual biomass pellets during the production process. Real-time measurement of the true density could be beneficial for the operation settings, such as the ratio of the binding agent to the raw material, the temperature of operation, the production rate, and the mixing ratio. The true density could also be used for rough measurement of the bulk density, which is a necessary parameter in commercial production. Therefore, knowledge of the true density is required during production in order to maintain the pellet quality as well as operation conditions. A prediction model was developed using partial least squares (PLS) regression across different wavelengths selected using different spectral pre-treatment methods and variable selection methods. After model development, the performance of the models was compared. The best model for predicting the true density of individual pellets was developed with first-derivative spectra (D1) and variables selected by the genetic algorithm (GA) method, and the number of variables was reduced from 256 to 53 wavelengths. The model gave R^2_{cal}, R^2_{val}, SEC, SEP, and RPD values of 0.88, 0.89, 0.08 g/cm^3, 0.07 g/cm^3, and 3.04, respectively. The optimal prediction model was applied to construct distribution maps of the true density of individual biomass pellets, with the level of the predicted values displayed in colour bars. This imaging technique could be used to check visually the true density of biomass pellets during the production process for warnings to quality control equipment.

1. Introduction

The use of biomass is increasing because it is a form of renewable energy that is broadly accepted as a substitute for fossil fuels (Saosee et al., 2020). Biomass is organic material used to generate heat energy. Biomass, such as wood chips, sawdust, and agricultural residues such as rice husks, rice straw, bagasse, sugarcane leaves, palm kernel shells, and coconut shells, is waste from agro-industrial factories. These waste products are processed into biomass pellets in order to increase their value. Nowadays, biomass in pellet form is favoured because it provides a low moisture content, high density, and consistent size. Biomass pellets are also more convenient to store, and they are easy to manage, move, and transport due to their grain-like flow characteristics (Zafar, 2022).

Forecasts of future global wood pellet demand predict a significant increase of up to 54 million tonnes in 2024 (40% demand for the heating pellet market and 60% demand for the industrial pellet market) (Strauss, 2020). Hence, a technology that helps to develop the production process to ensure its efficiency is required. In a commercial context, biomass pellets must meet the standard quality specifications. The heating value, ash, moisture content (MC), bulk density, and so on must meet the criteria of the standard quality specifications in order to trade the biomass pellets (García-Maraver et al., 2011). The quality of the biomass pellets is also an important factor in the combustion process for the industrial sector (Department of Alternative Energy Development and Efficiency, 2012). In general, the quality of biomass pellets is determined in the laboratory by using the standard ASTM E871–872 method (ASTM International, 1982). To satisfy the just-in-time production system, the quality of the biomass needs to be measured rapidly or in parallel during production. Therefore, real-time/in-line/at-line and rapid measurement of biomass pellet quality is necessary in the industry.

* Corresponding author at: Department of Agricultural Engineering, Faculty of Engineering, Khon Kaen University, Khon Kaen 40002, Thailand.
E-mail address: jetspo@kku.ac.th (J. Posom).

https://doi.org/10.1016/j.aiia.2022.11.004

According to previous studies related to the inspection of biomass pellets, near-infrared (NIR) spectroscopy techniques have been used for rapid and non-destructive testing. Posom and Sirisomboon (2015) predicted the MC and heating value of *Jatropha curcas* kernels using Fourier transform near-infrared (FT-NIR) spectroscopy (wavenumber range of 1,250,000–400,000 m^{-1}) (R^2 and RMSEP are 0.97 and 4.0%, respectively, for moisture content prediction; R^2 and RMSEP are 0.86 and 360 J/g, respectively, for heating value prediction). Posom and Sirisomboon (2017) estimated C, H, N, O, and S and the lower heating value of bamboo using FT-NIR spectrometry (12,500–3600 cm^{-1}) and reported RMSEP values of 0.532%, 0.0427%, 0.0276%, 1.110%, 0.00761%, and 0.119 MJ/kg, respectively. Gillespie et al. (2015) evaluated the MC, C, ash content, and gross calorific value of biomass pellets using NIR spectroscopy (880–1720 nm), finding RMSEP values of 0.503%, 0.089%, 0.837%, and 0.912 MJ/kg, respectively. Pitak et al. (2021a) applied NIR hyperspectral imaging of biomass pellets for the prediction of the fuel ratio (FR), volatile matter (VM), fixed carbon (FC), and ash content (A), as well as C, H, N, and S (Pitak et al., 2021b). Likewise, Feng et al. (2018) used this method to predict the moisture content (MC), ash content (A), volatile matter (VM), fixed carbon (FC), and calorific value (CV). As in previous works, they reported that NIR spectroscopy had a high potential for in-line measurement of biomass qualities.

Important parameters of biomass pellets are their bulk density and true density. Bulk density is defined as the mass of the total number of particles of the material divided by the total volume they occupy, while the true density of solids is the ratio of the mass of the solid to the volume occupied by that solid body (García-Maraver et al., 2011). Although the true density is not a quality criterion for biomass pellets, it can be checked rapidly and monitored during the production process to confirm the optimal ratio of the binding agent to the raw material, the initial moisture content used, the temperature of operation, the production rate, or operation conditions (da Silva et al., 2020). Moreover, the true density can be used to make a rough estimate of the bulk density (da Silva et al., 2020). The true density values of biomass pellets are always higher than the bulk density values because the volume of voids is excluded from the calculation, and bulk density has a correlation with the true density (Theerarattananoon et al., 2011).

Therefore, real-time measurement of the true density might be an innovative method for controlling biomass pellet production in the future. If the quality of the biomass pellets can be checked quickly and accurately, this will allow better quality control and make it easier to plan the production to keep up with the demands of the market.

NIR hyperspectral imaging has a high potential to be used for real-time measurement, and its main advantage is that it facilitates the visualization of the distribution of different chemical components in a sample (Manley, 2014).

However, too large data sets of hyperspectral images can be a barrier to analysis and in-line application. Therefore, reduction of the data set is necessary. Wavelength selection was therefore used to reduce the number of wavelengths. Using a few significant wavelengths in developing the calibration model may be at least as accurate as using a full range of wavelengths, and using just a handful of wavelengths is simple and cost-effective (Liu et al., 2014). Appropriate chemometric methods play an important role in the successful prediction of the density of individual pellets by NIR hyperspectral imaging. Therefore, we decided to use wavelength selection methods combined with a multivariate technique for developing the calibration model. Thus, this research aims 1) to investigate a suitable model for predicting the true density of individual biomass pellets using the NIR hyperspectral imaging technique, including the effect of variable selection and spectral pre-treatment, and 2) to create distribution maps of a pellet sample for in-line applications.

2. Materials and methods

2.1. Sample

A total of 140 biomass pellets were used in the experiment, obtained from 11 agricultural waste materials: filter cake (11 pellets), *Leucaena leucocephala* (9 pellets), bamboo (9 pellets), cassava rhizome (15 pellets), bagasse (14 pellets), sugarcane leaves (15 pellets), straw (15 pellets), rice husk (14 pellets), eucalyptus bark (11 pellets), Napier grass (13 pellets), and corn cob (14 pellets). All these agricultural waste materials were transformed into pellets using a palletization machine (KN-D-200, Tianjin Zhengyao Machinery, China). All the samples were then sealed in bags and stored at room temperature of about 25 ± 2 °C. Because the aim of this study was to estimate the true density of biomass pellets produced from the production process, standardized biomass pellet samples, with standardized parameters such as moisture content ≤10% and a pellet diameter of 6 mm, were used in the experiments to predict the true density. This condition facilitates the measurement of NIR hyperspectral imaging and sample collection.

2.2. NIR hyperspectral imaging

The NIR hyperspectral imaging system includes a hyperspectral camera (Imspector N17E, Specim, Finland), a CCD camera (Xeva 992, Xenics Infrared Solutions, Belgium), two 500 W tungsten halogen light sources (Lowel Light Inc., NY, USA), and control software (Specim's LUMO Software Suite, Spectral Imaging Ltd., Oulu, Finland), with 320 pixels (with a spatial resolution of 30 μm per pixel) in the x-direction and 497 pixels in the y-direction. The NIR hyperspectral imaging provided a wavelength range between 900 and 1700 nm, and the resolution was 3.2 nm. Approximately 11 pellets were placed on a sample holder made of acrylic sheets and placed on a conveyor belt, whose speed was 10 mm/s, for scanning. Before scanning, the camera was calibrated by 1) setting the frame rate (25 Hz) and exposure time (8 ms) of the camera and 2) calibrating the NIR hyperspectral camera by scanning dark and white reference, and then scanning pellet sample, respectively. After scanning, the relative reflectance (R) values were calculated from Eq. (1):

$$R = \frac{R_{sample} - R_{dark}}{R_{white} - R_{dark}} \quad (1)$$

where R_{sample} is the absolute reflectance value of the biomass pellets, R_{dark} is the reflectance of the background (dark reference), and R_{white} is the reflectance value of the acrylic sheet (white reference). The R_{dark} was collected when the light source unit was turned off, and the R_{white} was obtained by scanning a Teflon whiteboard. R was used to create a prediction model in the next step.

2.3. Reference methods

After scanning, the true density was determined using the standard ASTM E 873 method (ASTM International, 2013). The length of the biomass pellets was measured using a vernier caliper (Eagle One) of size 0–150 mm with a resolution of 0.05 mm, and then the pellets were weighed with digital scales (OHAUS PR224/E, maximum capacity 120 g, readability 0.1 mg). The true density was calculated using Eq. (2) (ASTM International, 2013):

$$\text{true density}\left(\frac{g}{cm^3}\right) = \frac{\text{mass (g)}}{\text{volume (cm}^3\text{)}} \quad (2)$$

After that, the moisture content was determined using oven drying, with samples dried at 105 °C until they reached a constant weight.

2.4. Model development

Fig. 1 shows the flow diagram for the process of model development. After scanning, the three-dimensional (3D) images of pellet samples were obtained. They were converted to relative reflectance data. The region of interest (ROI) of the 3D NIR image was extracted using principal component analysis (PCA). This process was carried out to distinguish between biomass pellet pixels and the background pixels. After that, the values of the background pixels were changed to zero, and the ROI pixels of each pellet were averaged to one spectrum. The averaged reflectance of each pellet was thus the representative spectrum and was applied for model development.

The optimal model for predicting the true density was investigated. The model was developed using different spectral pre-treatments, and the variables (relative reflectance value at its corresponding wavelengths) selected using different methods were compared. Before model creation, either the raw spectra or the pre-treated spectra (including the first derivative [D1], second derivative [D2], and standard normal variate [SNV]) were used for model development (Liu et al., 2014). The model was optimized using selected wavelengths obtained from the correlation method (CM), variance method (VM), covariance method (COVM), and successive projections algorithm (SPA) and genetic algorithm (GA) methods (Sratthaphut and Ruangwises, 2012; Fei and Yong, 2009). After that, their performance was compared.

The variable selection was investigated, including CM, VM, COVM, SPA, and GA. For variable selection by CM, the correlation (r) between X (spectral data) and Y (true density) was determined. The absolute values of r were determined, and its corresponding wavelengths were

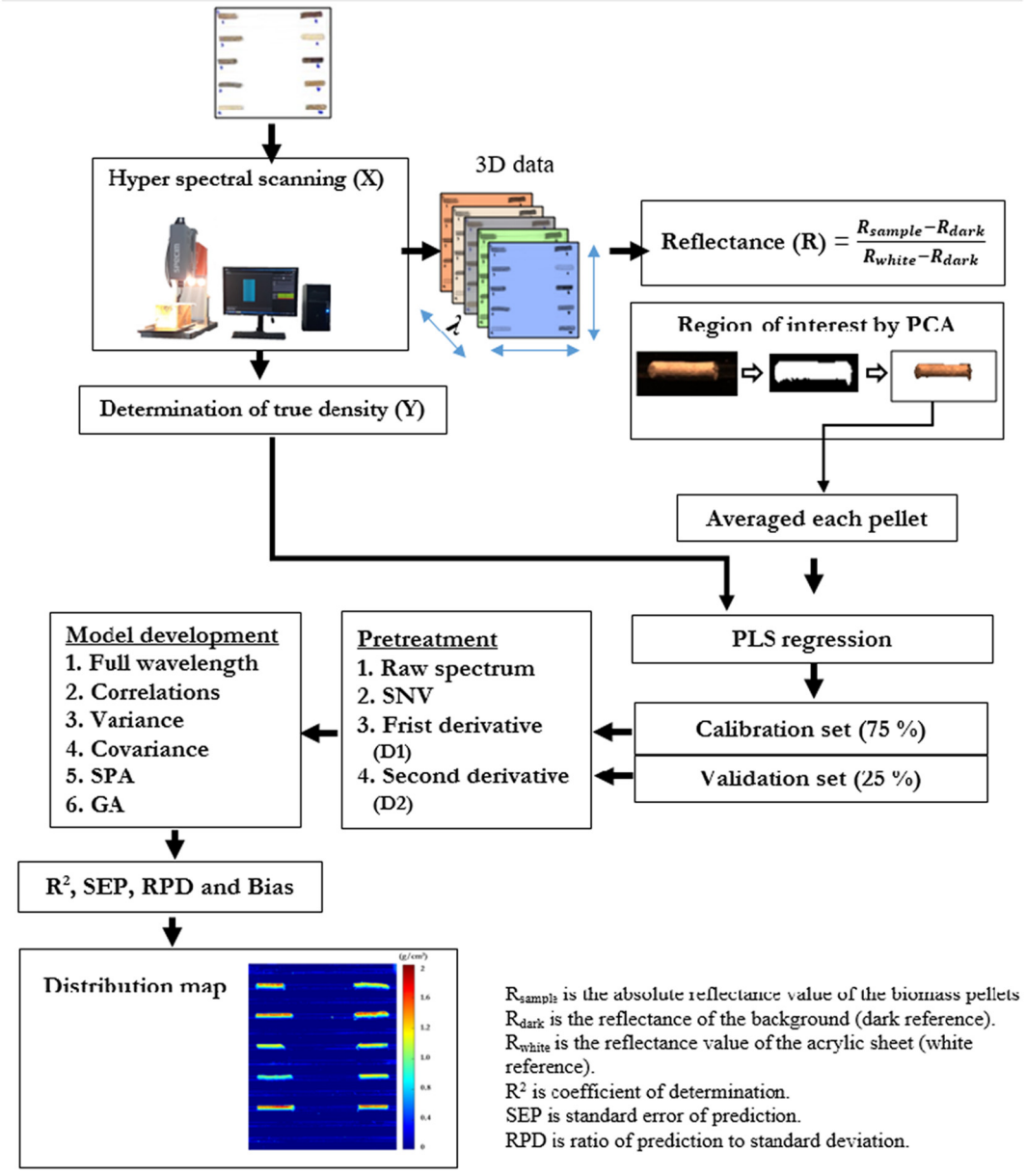


Fig. 1. Flow diagrams for prediction of true density (g/cm^3) by NIR hyperspectral imaging.
D1: 1st derivative spectra, (segment = 5), (gab = 5).
GA: genetic algorithms.
SPA: successive projections algorithm.

arranged in descending order. Then, the X data were arranged, with the wavelength having the highest |r| in the first column and the wavelength having a lower |r| in the next column. For VM, for every wavelength of X, the variance (var(X)) was determined. The X data were arranged in descending order according to the variance values of X, from the highest to the lowest value. Then, the arranged X data were collected. For COVM, the co-variances between Y and X_1, X_2, X_3, … X_n were computed, where n is the number of wavelengths. The X data were arranged in descending order according to the co-variance values of X and Y, from the highest to the lowest value, and then the arranged X data were obtained (Edelmann et al., 2021). The SPA method starts with one wavelength and then sequentially selects a new wavelength with the maximum projection value on the orthogonal subspace of the previously selected wavelength (Wang et al., 2016). The projections were carried out on the X matrix in order to display the least collinearity with the previous ones (Galvão et al., 2007). GAs are a type of mathematical model inspired by Charles Darwin's idea of natural selection, and they are another class of techniques used for wavelength selection in chemometrics. A GA allows a population composed of many individuals to evolve under specified selection rules to a state that maximizes the 'fitness' (Sratthaphut and Ruangwises, 2012). A combination of GA and partial least squares (PLS) regression has the advantages of both GAs (maximum fitness) and PLS regression (maximum covariance between the absorbance value and measured value) (Sratthaphut and Ruangwises, 2012). Pitak et al. (2021b) reported that the model for elemental prediction (C, H, N, and S) developed using the GA wavelength and PLS regression provided a higher accuracy than the model developed using the full wavelength range.

For the variable selection method, the most important wavelengths are arranged in the first column, and the next most significant wavelengths are arranged in the next column, up to the last wavelength. For model development, the total data were divided into two sets: the calibration set and the validation set. The data were analysed using the MATLAB program (R2019b, USA). The calibration model consists of matrices X (m × n) and Y (m × 1). Matrix X consists of the relative reflectance (R) of the pellets of the m samples recorded at n wavelengths. Matrix Y contains the reference data determined using the reference method (true density). The models were generated using Y and X_1, $X_{1\ to\ 2}$, $X_{1\ to\ 3}$, … $X_{1\ to\ n}$ using partial least squares regression (PLSR) and validated by one-fold cross-validation. After that, the number of wavelengths and the corresponding spectral pre-treatment method providing the lowest standard error of cross-validation (SECV) were selected to test the validation set; this procedure yields the optimal model.

The performance of the models was calculated based on statistical terms i.e. the coefficient of determination (R^2), standard error of calibration (SEC), standard error of prediction (SEP), ratio of prediction to standard deviation (RPD), relative standard error of prediction (RSEP), and bias:

$$\mathrm{SEC, SEP} = \sqrt{\frac{\sum_{1}^{m} (F - \bar{F})^2}{m}} \tag{3}$$

$$R^2 = 1 - \frac{\sum_{1}^{m} (F)^2}{\sum_{1}^{m} (Y - \bar{Y})^2} \tag{4}$$

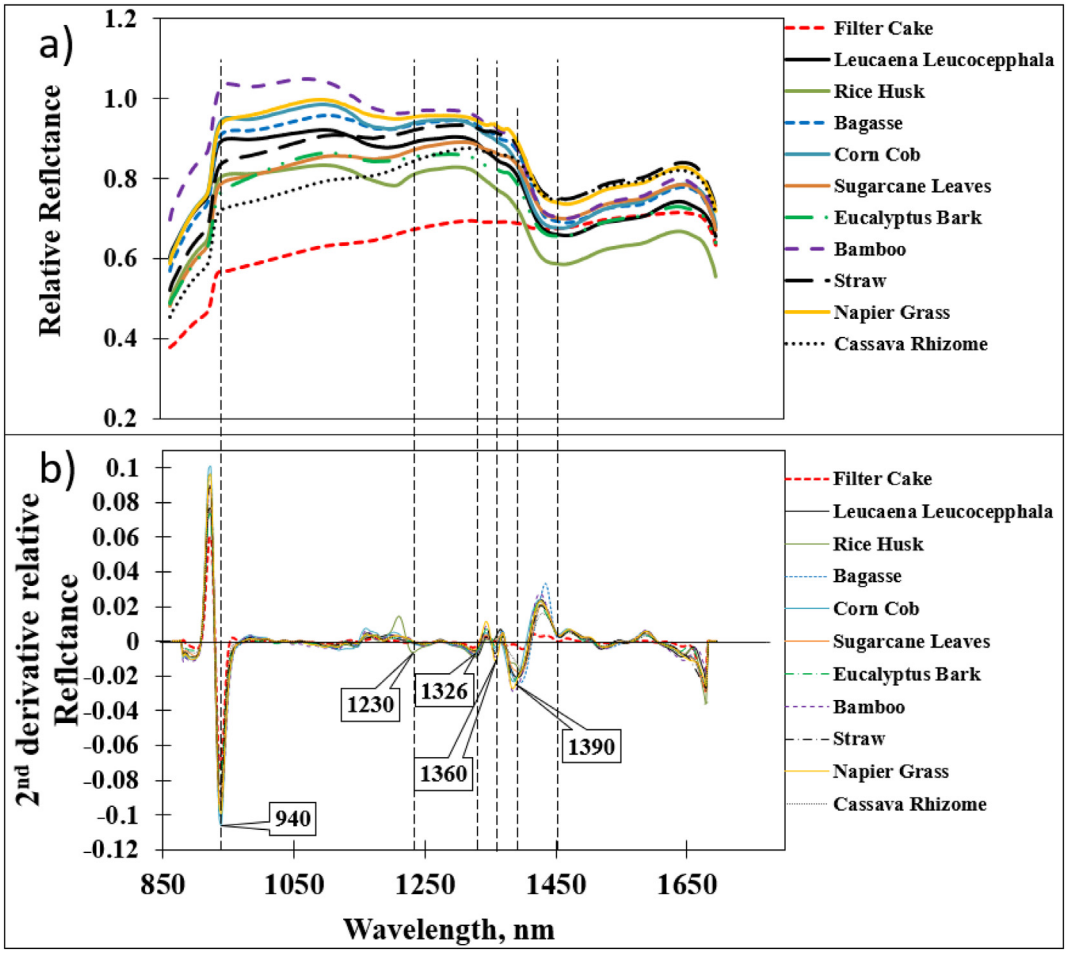


Fig. 2. Average spectra of 11 types of various biomass pellets: a) raw spectra and b) second derivative (D2).

Table 1
Statistical data of moisture content (%) and true density (SD, g/cm^3) of all biomass types obtained by ASTM E 873.

Parameter	n	Max	Min	Mean	Range	std
SD, g/cm^3	140	1.48	0.53	1.19	0.23	0.23
MC, %	140	10.04	2.36	7.60	7.68	1.77

std: standard deviation.

$$\text{Bias} = \frac{\sum_{1}^{m}(F)}{m} \quad (5)$$

$$\text{RPD} = \frac{SD}{SEP} \quad (6)$$

where m is the number of samples and $\overline{Y}$ is the mean value of all Y values.

The model performance was indicated by the criteria R^2 and RPD: $R^2 > 0.90$ and RPD > 3 indicates an excellent prediction model, $0.81 < R^2 < 0.90$ and $2.5 < RPD < 3$ indicates a good prediction model, $0.66 < R^2 < 0.80$ and $2.0 < RPD < 2.5$ indicates an approximate prediction model, and $R^2 < 0.66$ and RPD < 2 indicates a poor prediction model (Zornoza et al., 2008).

2.5. Visualization of the true density in the distribution map

The optimal model was obtained, including the optimal PLS factors, spectral pre-treatment method, and selected wavelength. These were applied to predict the true density and then used to create the spatial prediction map, which was the visualization of the individual true density. The spectrum in each pixel was used to calculate the predicted value (Ypre). The Ypre in each pixel channel was calculated from Ypre = X × B, where X is the relative reflectance in each pixel obtained from the test set, and B is the regression coefficient obtained from the optimal PLS regression. Ypre was represented as a two-dimensional (2D) array of pixels. The predicted density was represented using a colour level on a 2D distribution map (Su, 2020), which showed the spatial variations in the true density. The true density of each pellet was assessed from the distribution map,

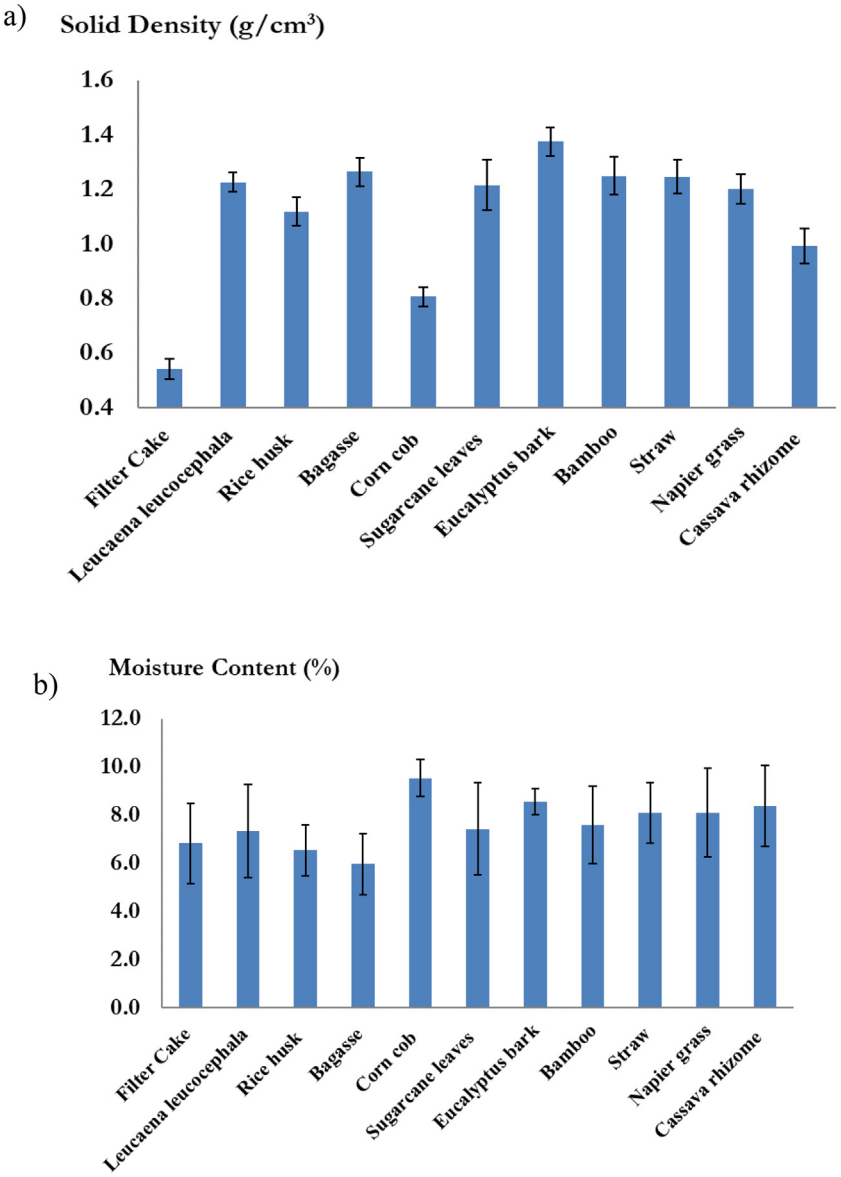


Fig. 3. The standard deviation of a) true density, and b) moisture content of the 11 varieties of biomass pellets.

which illustrated in different colors the predicted true density. The accuracy of the prediction was assessed by comparing, one by one, the acquired distribution maps and the pellets whose true density was known from the validation set. This information could be used for visualizing in-line measurements in quality assessment and control. NIR hyperspectral image analysis was carried out using MATLAB (R2019b, USA).

3. Results and discussion

3.1. NIR spectra of different biomass pellets

Fig. 2a and b show the averaged raw spectra and D2 spectra of the 11 biomass pellet types, respectively. The spectral profile of each biomass pellet is similar, but the relative reflectance values differ. While relative reflectance of fitter cake had a quite different compared to other biomass. It can be seen that low true density had also low reflectance value, while high true density had higher reflectance value. Solid density is the ratio of the measured mass to the measured volume. Low density had more air gaps in the material, which air can absorb light more than organic matter. It causes a pellet with a low density had a lower reflectance value. The obvious negative reflectance values of the biomass pellets of the D2 spectra were at the wavelengths of 940 nm, 1230 nm (vibrational band of C—H (3ν) in hydrocarbon), 1326 nm, 1360 nm (C—H combination in hydrocarbon), 1390 nm (C—H methyl in hydrocarbon) and1450 nm (O—H (2ν), .O-H in starch) (Williams, 2007; Burns and Ciurczak, 1992).

Yang et al. (2016) reported similar results. For wood pellets, there were many absorbance bands in the wavelength range of 1100–2500 nm, with the highest reflectance at the wavelengths of 1473, 1925, 2092, 2267, and 2328 nm. Gillespie et al. (2015) reported the wavelengths that impacted the prediction of MC, with an absorbance at the wavelengths of 1384 nm (1st overtone of the O—H bond structure) and 915 nm (3rd overtone of the C—H bond structure). Pitak et al. (2021b) reported a high reflectance value at the wavelengths of 938, 1121, 1248, 1323, 1384, and 1440 nm of the second derivative (D2) spectra.

Table 3
Result of model for predicting true density of various biomass pellets in the validation set based on different effects.

MC/true density	N	SEP (g/cm^3)	Bias (g/cm^3)
MC			
0–5%wt	4	0.089	−0.008
5–10%wt	30	0.074	0.003
Total	34		
Variety			
Filter cake	2	0.016	−0.051
Leucaena leucocepphala	2	0.149	0.058
Bamboo	2	0.037	0.065
Cassava rhizome	4	0.028	0.036
Bagasse	2	0.025	−0.099
Sugarcane leaves	5	0.049	0.060
Straw	3	0.087	0.004
Rice husk	2	0.007	−0.004
Eucalyptus bark	6	0.073	0.021
Napier grass	3	0.022	−0.087
Corn cob	3	0.026	−0.070
Total	34		

3.2. Reference value

Table 1 presents the statistical properties of the true density (SD, g/cm^3) and moisture content (MC, %) of the 11 biomass pellet types using the ASTM E 873 method, with SD values in the range of 0.53–1.48 g/cm^3 and MC in the range of 2.36–10.04%. In the EN 14961–2 and ISO 17225-2 standards, the MC of wood pellets is <10% (ASTM International, 2013; Gemco Energy Machinery Co., Ltd., 2022), which corresponds to the scope of this research's aim to use pellets with standardized parameters, such as moisture content ≤10%.

Fig. 3a and b show the standard deviation values of the 11 types of biomass pellets. The true density is in the range of 0.04–0.09 g/cm^3, and the standard deviation values of the moisture content are between 0.5% and 1.9%.

This result is consistent with Feng et al. (2018), who reported that the wood pellet moisture content ranged from 5.96% to 13.34%. In 2012,

Table 2
Result of model for predicting true density of various biomass pellets in the calibration set and the validation set.

Variable selection method	Pretreatment	Number of wavelengths	F	Calibration set			Validation set					
				N	R^2_{cal}	SEC	n	R^2_{val}	SEP	RSEP	RPD	BIAS
Full wavelength	Raw	256	10	106	0.93	0.06	34	0.90	0.07	6.10	3.20	−0.01
	SNV	256	12	106	0.96	0.05	34	0.88	0.08	6.87	2.85	−0.01
	D1	256	10	106	0.94	0.06	34	0.87	0.08	6.91	2.82	−0.01
	D2	256	10	106	0.94	0.06	34	0.85	0.09	7.59	2.56	0.00
CM	Raw	173	9	106	0.92	0.06	34	0.87	0.08	6.97	2.80	−0.01
	SNV	132	9	106	0.88	0.08	34	0.81	0.10	8.40	2.34	−0.01
	D1	70	11	106	0.88	0.08	34	0.78	0.11	9.16	2.16	0.02
	D2	50	8	106	0.76	0.11	34	0.75	0.11	9.70	2.02	0.01
VM	Raw	172	9	106	0.91	0.07	34	0.86	0.09	7.36	2.65	0.00
	SNV	130	10	106	0.90	0.07	34	0.82	0.10	8.21	2.37	0.00
	D1	110	10	106	0.92	0.07	34	0.90	0.07	6.16	3.16	0.00
	D2	100	9	106	0.90	0.07	34	0.88	0.08	6.80	2.87	−0.01
COVM	Raw	172	8	106	0.88	0.08	34	0.87	0.08	7.00	2.79	−0.01
	SNV	120	10	106	0.90	0.07	34	0.80	0.10	8.77	2.23	0.01
	D1	100	10	106	0.92	0.07	34	0.90	0.07	6.10	3.19	0.00
	D2	145	7	106	0.84	0.09	34	0.79	0.10	8.89	2.19	0.00
SPA	Raw	177	9	106	0.93	0.06	34	0.88	0.08	6.84	2.84	0.00
	SNV	145	10	106	0.91	0.07	34	0.84	0.09	7.70	2.53	0.00
	D1	75	10	106	0.93	0.06	34	0.87	0.08	7.11	2.74	0.00
	D2	90	8	106	0.85	0.09	34	0.87	0.08	7.12	2.77	0.01
GA	Raw	100	9	106	0.92	0.07	34	0.90	0.07	6.26	3.12	0.00
	SNV	35	11	106	0.91	0.07	34	0.81	0.10	8.42	2.32	−0.01
	D1	53	9	106	0.88	0.08	34	0.89	0.07	6.40	3.04	0.00
	D2	45	9	106	0.89	0.08	34	0.86	0.08	7.23	2.69	0.00

N: number of samples in calibration set, F: PLS factors, R^2_{cal}: coefficient of determination of calibration set, SEC (g/cm^3): standard error of calibration set, N: number of samples in validation set, R^2_{val}: coefficient of determination of validation set, SEP (g/cm^3): Standard error of prediction, RSEP (%) = relative standard error of prediction, RPD: ratio of prediction to deviation, CM: correlation method, VM: variance method, COVM: co-variance method, successive projections algorithm (SPA), genetic algorithm (GA).

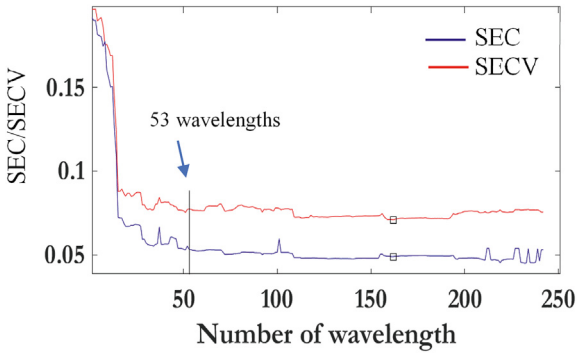


Fig. 4. SEC/SECV versus increasing wavelengths.

Samuelsson et al., 2012 reported the trend of density and moisture content using sawdust pellets and found that the moisture content affected the bulk density (kg/m^3). The moisture value was in the range of 8–14%, and the total density in the range of 500–750 kg/m^3. Moreover, da Silva et al. (2020) reported the relationship between mechanical durability (%) and the bulk density (kg/m^3), and they found that there was a correlation and a strong trend with $R^2 = 0.94$ and *P*-value = 0.004.

3.3. Results of the model development and the test-set method

Table 2 shows the results for predicting the true density of various biomass pellets developed from different spectral pre-treatments and different numbers of wavelengths from the respective variable selection methods, demonstrating the selection of the variables N, R^2_{cal}, SEC, R^2_{val}, SEP, RSEP, RPD, and bias. The prediction results depended on the development methods, spectral pre-treatment, and the number of wavelengths. Overall, the models with different variable selection methods gave the same performance, but the number of wavelengths differed.

The models developed with the full range of wavelengths (256 wavelengths)-raw spectra, VM (110 wavelengths)-D1 spectra, COVM (100 wavelengths)-D1 spectra, and GA (100 wavelengths)-raw spectra gave similar results, and the best accuracy for R^2 and SEP was 0.90 and 0.07 g/cm^3, respectively.

The method that had the fewest possible wavelengths, a low SEC, and a low SEP was the best method and could be selected as the prediction model. Hence, we recommend that the model developed with the D1 spectra and 53 wavelengths selected by the GA method was the optimal model, since R^2_{val}, SEP, RSEP, RPD, and bias were 0.89, 0.07 g/cm^3, 6.40%, 3.04, and 0.00 g/cm^3, respectively. The criterion for model selection was that if the performance was similar, the model developed with fewer wavelengths should be selected. This model indicates an excellent prediction model, could be used for quality control (Zornoza et al., 2008). Our result was more accurate than that of other NIR research studies. Sanseechan et al. (2018) used a portable NIR spectrometer for rapid screening of the true density of sugarcane stalk, and the best model in their study was developed with the pre-processing method of smoothing the moving average, and R^2, r^2, SEC, SEP, bias, and RSEP were 0.65, 0.67, 23.740 kg/m^3, 22.131 kg/m^3, −1.142 kg/m^3, and 2.000%, respectively. Li et al. (2022) investigated appropriate pre-processing, variable selection, and multivariate calibration techniques

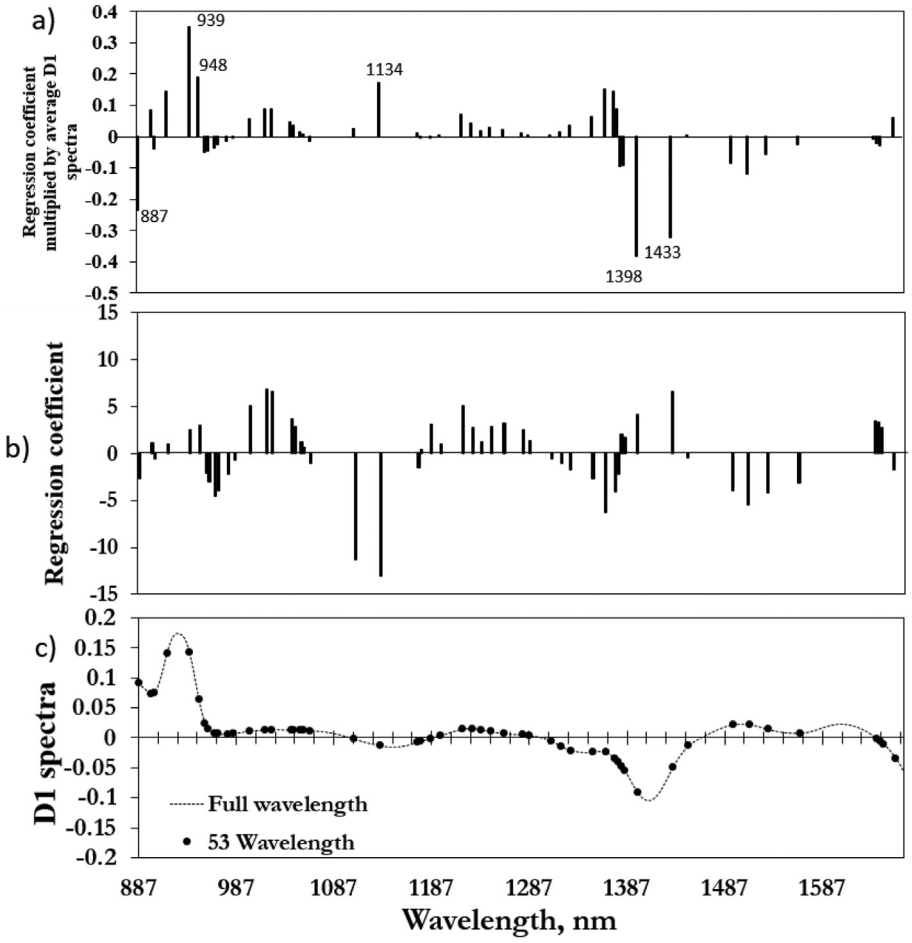


Fig. 5. Spectra and coefficients on model prediction; (a) multiplication of regression coefficient values and the average spectra of first derivative (D1) spectra; (b) regression coefficient from 53 selected wavelengths; and (c) D1 versus 53 wavelengths selected by GA.

to improve the density prediction accuracy of Chinese white poplar (*Populus tomentosa* Carrière) wood. The results showed that the best prediction was obtained by generalized regression neural network (GRNN) models combined with the lifting wavelet transform (LWT) and the competitive adaptive reweighted sampling (CARS) method ($R^2p = 0.870$, RMSEP = 13 kg/m^3, RPDp = 2.774).

Table 3 shows the result of the model for predicting the true density of the various biomass pellets in the validation set based on different effects. There are fewer samples in the moisture content range of 0–5% than in the range of 5–10%. For the prediction of the true density in the moisture content range of 5–10%, the SEP and bias were 0.074 and 0.003 g/cm^3, respectively. The variety effect found that the SEP and bias of rice husk were the lowest, with an SEP and bias of 0.007 and − 0.004 g/cm^3, respectively.

Therefore, this model was excellent and could be used for quality measurement. The RSEP (relative standard deviation of prediction), which is the relative ratio of the absolute error of prediction to a reference value, showed that every predicted value gave an error of approximately 6.40% of the reference value (John et al., 2007).

Fig. 4 shows SEC/SECV versus the number of wavelengths developed with the GA-D1 method. The 53-wavelength model was selected because the SEC and SECV remained constant and because increasing the number of wavelengths did not increase the accuracy.

Fig. 5 shows spectra and coefficients on model prediction. Fig. 5a illustrates the values obtained by the multiplication of regression coefficient values (see Fig. 5b) and the average spectra of D1 derivatives spectra selected by the GA method (see Fig. 5c) as a function of wavelength. The predicted true density value was calculated with the sum of the multiplication of regression coefficient values and the average D1 spectra. Hence high-intensity peaks at any wavelength mean that those wavelengths in Fig. 5a strongly affect the model prediction.

According to the peaks in Fig. 5a, some obvious peaks with high values were at approximately 887, 939, 948, 1134, 1398 and 1433 nm, means that these wavelengths influence the prediction.

Fig. 6a and b also show the results of the predicted and measured values of the calibration set and validation set, considering two different moisture ranges: 0–5% and 5–10%. The result shows that the different MC ranges do not affect the prediction. Fig. 6c and d show the scatter plots of the measured and predicted values of the calibration set, and the validation set for the true density is affected by the different biomass pellet types, which do not affect the accuracy.

3.4. Visualization of the true density results in a distribution map

Fig. 7a shows the RBG image of the biomass pellets. Fig. 7b shows the distribution map of the quality level of each pellet in the different

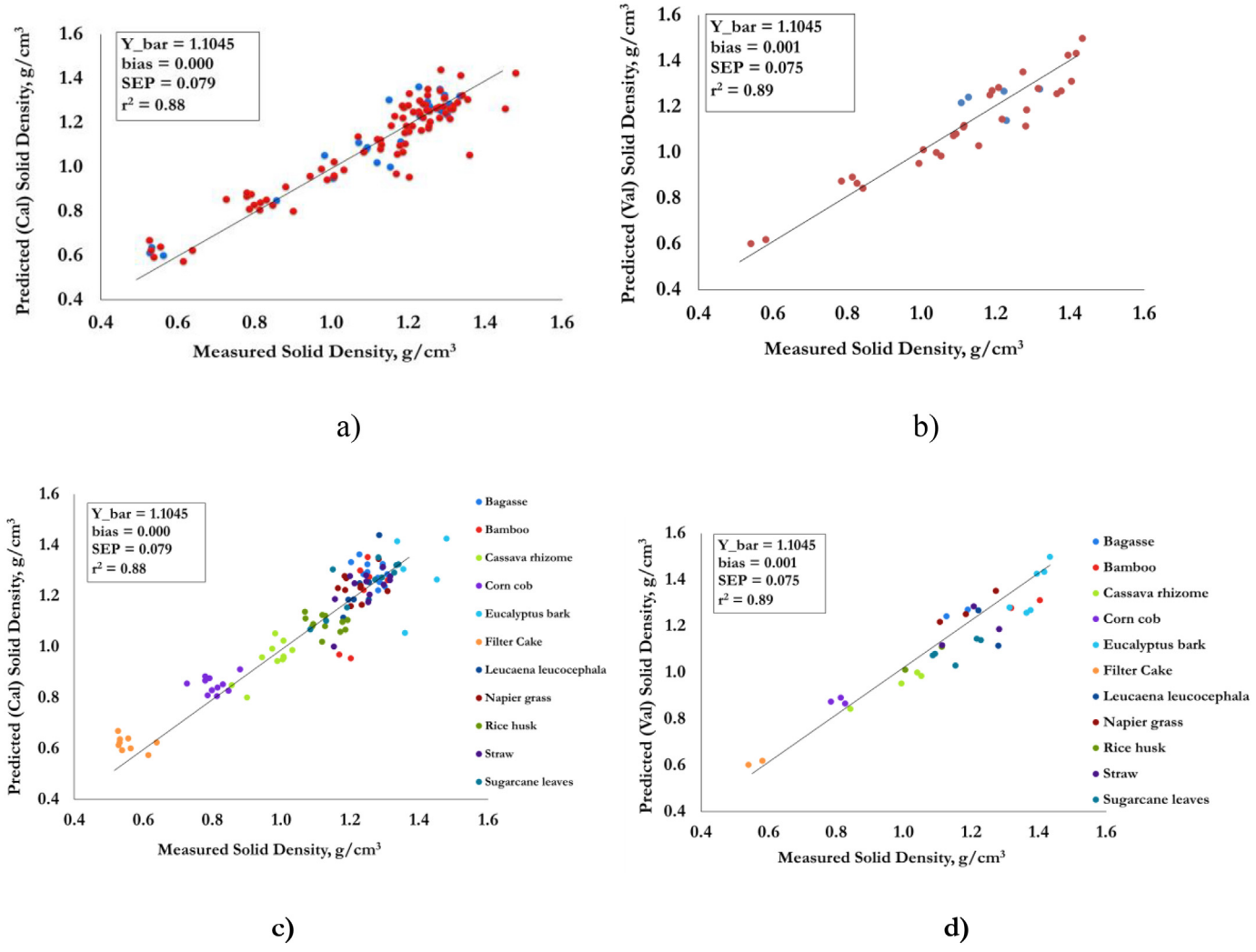


Fig. 6. Scatter plots between measured and predicted value in the prediction of true density of a) calibration set, b) validation set affected by moisture content (MC range 0 to 5% = red colour points and > 5 to 10% = blue points), and c) calibration set, d) validation set affected by variety.

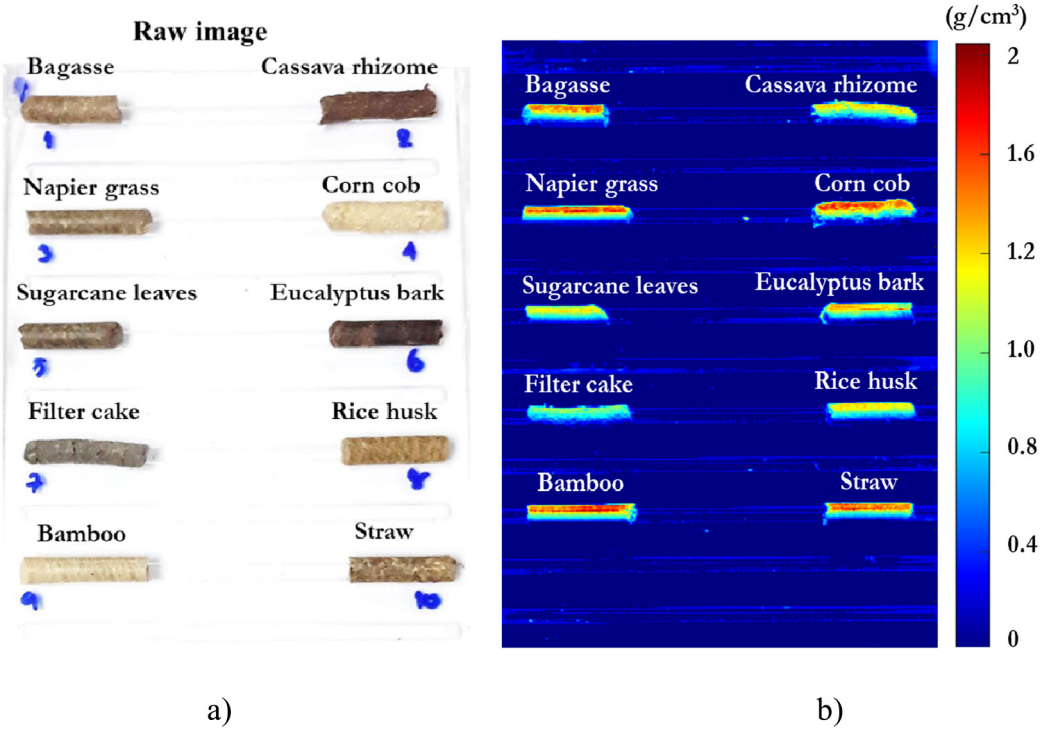


Fig. 7. a) RGB image, b) predictive distribution maps of quality of different biomass pellet types.

biomass pellet types predicted by the D1 spectra coupled with 53 wavelengths from the variable selection method of the GA model. The predicted value of each pixel was then converted into a different colour scale. The predictive distribution maps were converted into 2D NIR images of the true density. The highest pellet density values are shown in red, medium values are shown in green, and the lowest value is shown in blue. Other colour of predictive image in the linear colour scale changed according to the level of predicted true density. The predicted image demonstrated irregular colors in each pixel of each pellet because true density in each pixel of each pellet was non-homogeneous. Other causes, the spectrum in each pixel obtained from different angles of incidence and reflection were different value. Even though the same pellet, the reflectance values in each pixel were not equal. However, overall colors of the pellet images from different true density were clearly different. Therefore, the true density of pellet could be predicted by observing the colors of the images. The higher true density contained more red while the lower true density contained more blue. Displaying colour images of the properties of biomass pellets may be convenient and easy to apply online in the industry.

4. Conclusions

The NIR hyperspectral imaging technique is proposed for rapid measurement of true density of biomass pellets. The prediction model were developed using the full wavelength range, CM, *VM*, COVM, SPA, and GA coupled with four pre-treatment methods including raw spectra, D1, D2, and SNV. Then, their performances were compared. The result demonstrated that true density model was excellent and could be used for quality control. The most appropriate method predicts the true density of GA biomass pellets and spectral conditioning with D1 by using 53 wavelengths. Then, a distribution map can show the specific properties of the biomass pellets, in a form that can be easily viewed and used in various procedures in the industrial sector, such as quality control in the production process and quality selection before the product is sold. The model can also be developed for other materials and applications in the future. The appropriate model and important wavelength could be used for creating a low-cost camera for real time measuring of true density upon conveyor belt for further.

CRediT authorship contribution statement

Lakkana Pitak: Data curation, Writing – original draft, Conceptualization, Methodology, Software. **Khwantri Saengprachatanarug:** Validation. **Kittipong Laloon:** Validation. **Jetsada Posom:** Supervision, Conceptualization, Methodology, Writing – review & editing, Validation, Funding acquisition, Project administration.

Declaration of Competing Interest

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

Acknowledgements

This research project is supported by Research and Graduate Studies, Khon Kaen University, Khon Kaen, Thailand.

References

ASTM International, 2013. Standard Test Method for Bulk Density of Densified Particulate Biomass Fuels. Published June https://doi.org/10.1520/E0873-82R13.

ASTM International. Standard Method for Moisture Analysis of Particulate Wood Fuels (ASTM E871-872), 1982. . Published DecemberAnnual Book of ASTM Standardsvol. 05.05.

Burns, D.A., Ciurczak, E.W., 1992. In: Burns, Donald A., Ciurczak, Emil W. (Eds.), Handbook of near-Infrared Analysis. 13. Marcel Dekker, INC, America. https://doi.org/10.1201/9781420007374.

da Silva, S.B., Arantes, M.D.C., de Andrade, J.K.B., Andrade, C.R., de Cassia Oliveira Carneiro, A., de Paula Protasio, T., 2020. Influence of physical and chemical compositions on the properties and energy use of lignocellulosic biomass pellets in Brazil. Renew. Energy 147, 1870–1879. https://doi.org/10.1016/j.renene.2019.09.131.

Department of Alternative Energy Development and Efficiency, 2012. Study Project of Biomass Pellet Standard for Development as a Biomass Fuel in the Future. Final Report; Faculty of Engineering and Industrial Technology Silpakorn University, Nakhon Pathom, Thailand, pp. 1–272.

Edelmann, D., Móri, T.F., Székely, G.J., 2021. On relationships between the Pearson and the distance correlation coefficients. Stat. Probab. Lett. 169, 108960. https://doi.org/10.1016/j.spl.2020.108960.

Fei, L., Yong, H., 2009. Application of successive projections algorithm for variable selection to determine organic acids of plum vinegar. Food Chem. 115, 1430–1436. https://doi.org/10.1016/j.foodchem.2009.01.073.

Feng, X., Yu, C., Shu, Z., Liu, X., Yan, W., Zheng, Q., Sheng, K., He, Y., 2018. Rapid and non-destructive measurement of biofuel pellet quality indices based on two-dimensional near infrared spectroscopic imaging. Fuel 228, 197–205. https://doi.org/10.1016/j.fuel.2018.04.149.

Galvão, R.K.H., Araújo, M.C.U., Silva, E.C., José, G.E., Soares, S.F.C., Paiva, H.M., 2007. Cross-validation for the selection of spectral variables using the successive projections algorithm. J. Braz. Chem. Soc. 18 (8), 1580–1584. https://doi.org/10.1590/S0103-50532007000800021.

García-Maraver, A., Popov, V., Zamorano, M., 2011. A review of European standards for pellet quality. Renew. Energy 36, 3537–3540. https://doi.org/10.1016/j.renene.2011.05.013.

Gemco Energy Machinery Co.,Ltd, 2022. Wood Pellets Quality Standards Review. http://www.biofuelmachines.com/wood-pellets-quality-standards-study.html [accessed 1 March 2022].

Gillespie, G.D., Everard, C.D., Mcdonnell, K.P., 2015. Prediction of biomass pellet quality indices using near infrared spectroscopy. Energy 80, 582–588. https://doi.org/10.1016/j.energy.2014.12.014.

John, S.S., Workman, J.W. Jerome Jr, Mark, O.W., 2007. NIRS applications in agricultural products. Handbook of Near-Infrared Analysis, 3rd ed. CRC press, pp. 347–386.

Li, Y., Wang, G., Guo, G., Li, Ya., Via, B.K., Pei, Z., 2022. Spectral pre-processing and multivariate calibration methods for the prediction of wood density in Chinese white poplar by visible and near infrared spectroscopy. Forests 13, 62. https://doi.org/10.3390/f13010062.

Liu, D., Sun, D.W., Zeng, X.A., 2014. Recent advances in wavelength selection techniques for hyperspectral image processing in the food industry. Food Bioprocess Technol. 7, 307–323. https://doi.org/10.1007/s11947-013-1193-6.

Manley, M., 2014. Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials. Chem. Soc. Rev. 43, 8200–8214. https://doi.org/10.1039/c4cs00062e.

Pitak, L., Laloon, K., Wongpichet, S., Sirisomboon, P., Posom, J., 2021a. Machine learning-based prediction of selected parameters of commercial biomass pellets using line scan near infrared-hyperspectral image. Processes 9 (2). https://doi.org/10.3390/pr9020316.

Pitak, L., Sirisomboon, P., Saengprachatanarug, K., Wongpichet, S., Posom, J., 2021b. Rapid elemental composition measurement of commercial pellets using line-scan hyperspectral imaging analysis. Energy 220, 119698. https://doi.org/10.1016/j.energy.2020.119698.

Posom, J., Sirisomboon, P., 2015. Evaluation of the moisture content of Jatropha curcas kernels and the heating value of the oil-extracted residue using near-infrared spectroscopy. Biosyst. Eng. 130, 52–59. https://doi.org/10.1016/j.biosystemseng.2014.12.003.

Posom, J., Sirisomboon, P., 2017. Evaluation of lower heating value and elemental composition of bamboo using near infrared spectroscopy. Energy 121, 147–158. https://doi.org/10.1016/j.energy.2017.01.020.

Samuelsson, R., Larsson, S.H., Thyrel, M., Lestander, T.A., 2012. Moisture content and storage time influence the binding mechanisms in biofuel wood pellets. Appl. Energy 99, 109–115. https://doi.org/10.1016/j.apenergy.2012.05.004.

Sanseechan, P., Panduangnate, L., Saengprachatanarug, K., Wongpichet, S., Taira, E., Posom, J., 2018. A portable near infrared spectrometer as a non-destructive tool for rapid screening of true density stalk in a sugarcane breeding program. Sens. Bio-Sens. Res. 20, 34–40. https://doi.org/10.1016/j.sbsr.2018.07.001.

Saosee, P., Shabbir, B.S., Gheewala, H., 2020. Feedstock security analysis for wood pellet production in Thailand. Energies 13, 5126. https://doi.org/10.3390/en13195126.

Sratthaphut, L., Ruangwises, N., 2012. Genetic algorithms-based approach for wavelength selection in spectrophotometric determination of vitamin B12 in pharmaceutical tablets by partial least-squares. Proc. Eng. 32, 225–231. https://doi.org/10.1016/j.proeng.2012.01.1261.

Strauss, W., 2020. Global Pellet Markets Outlook. https://www.canadianbiomassmagazine.ca/2020-global-pellet-markets-outlook/ (accessed 25 January 2020).

Su, W.H., 2020. Advanced machine learning in point spectroscopy, RGB- and hyperspectral-imaging for automatic discriminations of crops and weeds: a review. Smart Cities 3, 767–792. https://doi.org/10.3390/smartcities3030039.

Theerarattananoon, K., Xu, F., Wilson, J., Ballard, R., Mckinney, L., Staggenborg, S., Vadlani, P., Pei, Z.J., Wang, D., 2011. Physical properties of pellets made from sorghum stalk, corn stover, wheat straw, and big bluestem. Ind. Crop. Prod. 33 (2), 325–332. https://doi.org/10.1016/j.indcrop.2010.11.014.

Wang, J., Shi, T., Liu, H., Wu, G., 2016. Successive projections algorithm-based threeband vegetation index for foliar phosphorus estimation. Ecol. Indic. 67, 12–20. https://doi.org/10.1016/j.ecolind.2016.02.033.

Williams, P., 2007. Near-Infrared Technology Getting the Best out of Light Edition 5.0. A Short Course in the Practical Implementation of Near-Infrared Spectroscopy for the User. PDK Grain, Nanaimo, Canada.

Yang, Z., Li, K., Zhang, M., Xin, D., Zhang, J., 2016. Rapid determination of chemical composition and classification of bamboo fractions using visible near infrared spectroscopy coupled with multivariate data analysis. Biotechnol. Biofuels 9 (35). https://doi.org/10.1186/s13068-016-0443-z.

Zafar, S., 2022. Biomass Pelletization Process. https://www.bioenergyconsult.com/biomass-pelletization/ (accessed 1 March 2022).

Zornoza, R., Guerrero, C., Mataix-Solera, J., Scow, K.M., Arcenegui, V., Beneyto, J.M., 2008. Near infrared spectroscopy for determination of various physical, chemical and biochemical properties in Mediterranean soils. Soil Biol. Biochem. 40 (7), 30–1923. https://doi.org/10.1016/j.soilbio.2008.04.003.