JTEP Jurnal Keteknikan Pertanian,
Vol. 13 No. 2, p 318-339, 2025
P-ISSN 2407-0475 E-ISSN 2338-8439

Available online:
http://journal.ipb.ac.id/index.php/jtep
DOI: 10.19028/jtep.013.2.317-339

Prediction of Chemical Content in Robusta Coffee Beans Using Near Infrared Spectroscopy and Artificial Neural Network

Chorida Okasyari[1], I Wayan Budiastra[1,2*], Sutrisno[1,2]

[1]Department of Mechanical and Biosystem Engineering, Faculty of Agricultural Engineering and Technology, IPB University, Lingkar Akademik Street, IPB Dramaga Campus, Babakan, Bogor, West Java 16680, Indonesia.
[2]Center for Research on Engineering Application in Tropical Agriculture (CREATA), IPB University, Lingkar Akademik Street, IPB Dramaga Campus, Babakan, Bogor, West Java 16680, Indonesia.
*Corresponding author, email: wbudiastra@apps.ipb.ac.id

Article Info

Submitted: 21 May 2025
Revised: 8 July 2025
Accepted: 9 July 2025
Available online: 11 July 2025
Published: June 2025

Keywords:
calibration, coffee, NIRS, PCA-ANN, validation

How to cite:
Okasyari, C., Budiastra, I W., Sutrisno. (2025). Prediction of Chemical Content in Robusta Coffee Beans Using Near Infrared Spectroscopy and Artificial Neural Network. Jurnal Keteknikan Pertanian, 13(2): 318-339. https://doi.org/10.19028/jtep.013.2.318-339.

Abstract

The development of a rapid and cost-effective method to predict the chemical content of Robusta coffee is essential, considering that conventional chemical analyses are time-consuming and expensive. This study aimed to predict the moisture, protein, fat, ash, and carbohydrate contents of Pagar Alam Robusta coffee beans using near infrared spectroscopy (NIRS) combined with principal component analysis-artificial neural network (PCA-ANN). Reflectance spectra of the beans were measured using a NIRFlex N500 Spectrometer (1000–2500 nm), followed by chemical composition analysis using standard laboratory methods. Spectral data were pre-processed, then calibrated and validated against the reference using PCA-ANN. PCA-ANN with No1SG1 pre-treatment yielded optimal predictions for most chemical component of robusta coffe beans, with moisture content resulting in r =0.95 and RPD = 2.95, protein r = 0.92 and RPD = 2.38, fat r = 0.93 and RPD =2.19, and carbohydrate r = 0.93 and RPD = 2.29, while SG1 pre-treatment was more suitable for predicting ash content with r = 0.93 and RPD = 2.73. The results demonstrate that NIRS, combined with PCA-ANN, can accurately predict the chemical composition of Robusta coffee beans, offering a reliable alternative to conventional methods.

Doi: https://doi.org/10.19028/jtep.013.2.318-339

1. Introduction

Coffee is one of the main agricultural commodities that is widely cultivated in Indonesia. Based on 2023 data, the total national coffee production reached 758,725 tons, with the robusta variety contributing a dominant share of approximately 79.10% of the total production (BPS, 2024; Pusat Data dan Sistem Informasi, 2023). Pagar Alam is one of the cities in South Sumatra province that produces a large amount of robusta coffee (Irmeilyana et al., 2021). To date, coffee quality assessment in Pagar Alam relies only on sensory analysis, which produces highly variable assessment results. Quality assessment of robusta coffee beans can also be conducted using standard laboratory methods, but this method is time-consuming, requires complex sample preparation, and incurs high testing costs per sample.

As a solution to these issues, near infrared spectroscopy (NIRS) can be used as an alternative method because it can measure the chemical content of coffee beans non-destructively (Irawan et al., 2022). The ability to rapidly and accurately estimate chemical content without the need for special sample preparation makes NIRS highly potential applied (Osborne & Feran, 1993). Most studies have demonstrated the effectiveness of NIRS for the analysis of Arabica coffee. For example, Zhu et al. (2021) used a Nicolet 5700 FT-IR spectrometer to analyse green Arabica coffee and achieved a coefficient of determination (R2) = 0.98 (lipids) and R2 = 0.99 (protein).

In Indonesia, the development of NIRS for coffee analysis has also progressed. Rosita et al. (2016) predicted caffeine content in Gayo Arabica coffee beans using a NIRFlex N500 spectrometer with the partial least square (PLS) method, yielding r = 0.97. Yuwita et al. (2023) employed an NIRS FT-IR IPTEK T-1516 and PLS to predict the chemical content of Arabica coffee, achieving R2 = 0.99 for both caffeine and protein. These findings confirmed that NIRS is a reliable method for determining the chemical composition of Arabica coffee.

However, most of the NIRS studies mentioned previously still focused on Arabica coffee. A study conducted by Irawan et al. (2022) used the Thermo Nicolet Antaris II TM NIRS instrument on Robusta and Arabica coffee, applying the PLS method to predict the chemical content, resulting in r = 0.65 (moisture). Meanwhile, Setyawan et al. (2025) conducted a study on Robusta coffee using the NIRFlex N500 instrument and the PLS method, in which spectral transformation in absorbance form yielded the highest R^2 = 0.89 (moisture) and R^2 = 0.91 (lipid). Nevertheless, studies combining NIRS with an artificial neural network (ANN), a method known for its ability to handle nonlinear data relationships, remain limited.

Although linear methods such as PLS have shown good performance on Arabica coffee, the spectral characteristics and chemical composition of Robusta coffee are more complex due to higher protein content, as well as greater influence from growing location, altitude, variety, and postharvest processing. These factors cause variations in the NIRS spectral absorption patterns, which do not always exhibit linear relationships. Therefore, a nonlinear approach, such as an ANN, is required to better capture these complex patterns. ANN is capable of modelling nonlinear relationships through learning and training processes, potentially providing more accurate predictions (Alamsyah et al., 2021).

Aditama (2019) reported that combining NIRS with a principal component analysis-artificial neural network (PCA-ANN) successfully predicted the proximate and caffeine content in Arabica coffee, outperforming conventional methods such as PLS. However, Arabica and Robusta coffees differ significantly in their physical and chemical properties. Robusta contains more protein (18.5 g/100 g) compared to Arabica (14.3 g/100 g), and Robusta has a lower fat content (10.1%) compared to Arabica (16%) (Liu et al., 2022; Olango et al., 2025). These differences resulted in distinct NIRS spectral characteristics, suggesting that the PCA-ANN models developed for Arabica cannot be directly applied to Robusta without modification. Furthermore, a literature review found no previous studies applying NIRS and PCA-ANN approaches to predict the chemical composition of Robusta coffee, underscoring the novelty and significance of this study.

This study aimed to predict the moisture, protein, fat, ash, and carbohydrate contents of Robusta coffee beans from Pagar Alam using NIRS combined with PCA–ANN calibration, thereby addressing this gap and supporting the advancement of rapid, non-destructive analytical methods tailored to Robusta coffee.

2. Material and Methods

The methodological steps included spectral acquisition, chemical reference analysis, spectral pre-treatment, and calibration-validation using PCA-ANN and PLS models, as illustrated in Figure 1.

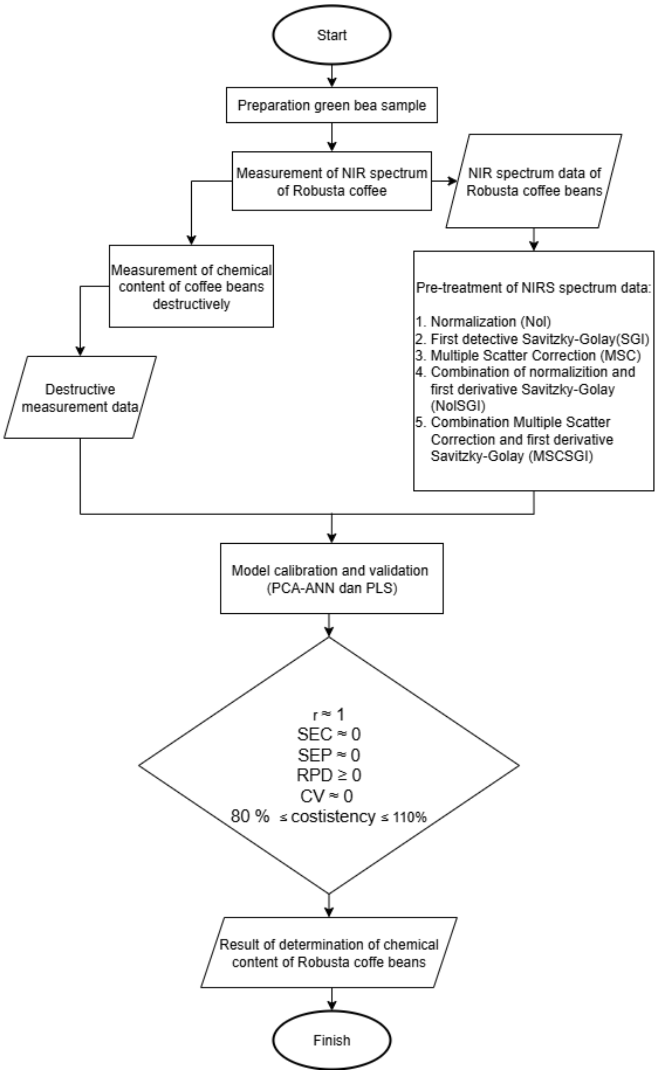


Figure 1. Flowchart of the Research Methodology.

This study used green robusta coffee beans (varieties: Rainbow and Red Pick) sourced directly from coffee farmers in Pagar Alam, South Sumatra Province. Sixty samples were collected. To anticipate overfitting, a value-based systematic hold-out validation method was used in which 60 samples were divided into two groups: a calibration set (two-thirds, 40 samples) and a validation set (one-third, 20 samples). The primary analytical instrument was a NIRFlex N500 spectrometer (BUCHI Labortechnik AG, Switzerland), operating in the wavelength range of 1000-2500 nm. Data processing and modeling were conducted using Unscrambler X 10.4, Microsoft Excel, and RapidMiner Studio 9.10 software.

2.1 Reflectance Measurement of Robusta Coffee Beans

The coffee bean samples were evenly and fully placed in petri dishes (66–73 grams) with a diameter of 10 cm and a height of 1.5 cm. Reflectance measurements were performed using the NIRFlex 500 spectrometer in the wavelength range of 1000-2500 nm, with the instrument set up as follows: data acquisition speed of 3 scans/s and operating temperature of 22–25°C. Each sample was measured three times, and the average value was calculated.

2.2 Destructive Chemical Content Measurement of Robusta Coffee Beans

The chemical components to be determined included moisture, protein, fat, ash, and carbohydrates. The moisture content was determined using the gravimetric method, as shown in Equation 1. Protein content was determined using the Kjeldahl method, as shown in Equation 2 (AOAC 2006). Fat content was measured using the Soxhlet extraction method, as shown in Equation 3 (AOAC, 2006). Ash content was determined using the gravimetric method, as shown in Equation 4 (AOAC, 2006). Carbohydrate content was calculated by the difference method using Equation 5, which involves subtracting the sum of other chemical components from 100 (AOAC, 2006).

$$\text{Moisture content (\%)} = \frac{(\mathrm{W3} - \mathrm{W1})}{(\mathrm{W2} - \mathrm{W1})} \times 100\% \qquad (1)$$

Where: W1= weight of the empty crucible (grams), W2= weight of crucible and sample before drying (grams), W3= weight of crucible and sample after oven drying (grams).

$$\%N = \frac{(mL\ HCLsample - mL\ HCLblank) \times NHCL \times 14{,}007}{sample\ weight\ (mg)} \times 100\%$$

Protein content (%) = % N × 6.25 (2)

$$\text{Fat content (\%)} = \frac{W3 - W1}{W2} \times 100\% \qquad (3)$$

Where: W1= weight of the empty fat flask (grams), W2= weight of the sample used for extraction (grams), Ws = weight of the fat flask and extracted fat after drying (grams).

$$\text{Ash content }(\%) = \frac{W3 - W1}{W2 - W1} \times 100 \tag{4}$$

Where: W1= weight of the empty crucible (grams), W2= weight of the crucible and sample before ignition (grams), W3= weight of the crucible and ash after ignition (grams).

2.3 NIRS Data Processing

Generally, applying spectral pre-treatment improves the calibration and validation results compared to spectra without pre-treatment (Budiastra et al., 2024). In this study, pre-treatment was performed using Unscrambler X 10.4 software and applied to the original spectra obtained from NIRS measurements. The process begins by importing the spectral data into the software, followed by defining the calibration and validation sets using the defined range function. Next, the transform option was selected from the task menu to apply the pre-treatment. Five spectral pre-treatment methods were employed in this study.

The first pre-treatment applied was normalization (No1), which was used to reduce the signal variation caused by non-chemical factors. The first derivative Savitzky–Golay (SG1) was applied to reduce overlapping chemical components (Bou-Orm et al., 2020). Third, multiplicative scatter correction (MSC) was used to correct the scattering effects caused by the physical differences among the samples. Fourth, a combination of No1 and SG1 (No1SG1) was applied to combine the benefits of normalization and SG1. Finally, a combination of MSC and SG1 (MSCSG1) was used to produce cleaner, more stable, and more representative spectral data for further analysis. The selection of pre-treatment was based on a previous study conducted by Aditama et al. (2019), which demonstrated that the configuration produced optimal model performance.

2.4 Model Calibration and Validation

The pre-treatment data were subsequently processed using principal component analysis (PCA) to reduce the dimensionality of the complex spectral data (Masithoh et al., 2021). PCA classifies spectral data into new variables called principal components (PC) (Dharmawan et al., 2023). These PC were then used as the input data for the ANN model. Using a PC as an input can enhance the performance of the model (Perera et al., 2021).

The ANN model applied in this study was a feed-forward neural network with a multilayer perceptron algorithm. The dataset was split into two-thirds for calibration and one-third for validation (Lengkey et al., 2013). Using a value-based systematic hold-out approach, the data were first sorted in descending order, and then alternately assigned to two groups. This approach was applied to ensure that the standard deviation values in both the calibration and validation sets were balanced, thereby minimizing the risk of overfitting.

The ANN architecture consists of three layers: the input, hidden, and output layers. Using fewer PC simplifies the analytical process (Aditama, 2019). This study used 5, 8, and 10 PCs based on an exploratory strategy to evaluate the effect of the number of PCs on the ANN model performance. The number of retained PCs represents different levels of cumulative variance, namely 5 PCs (low, ~80%), 8 PCs (medium, ~90%), and 10 PCs (high, > 95%) based on the results of the PCA, in order to assess how the number of PCs influences the prediction accuracy. The ANN architecture uses a single hidden layer with five neurons selected through trial and error to achieve optimal performance. The ANN output consisted of the predicted values for each chemical component (moisture, protein, fat, ash, and carbohydrates).

ANN was selected because of its ability to model complex and nonlinear relationships between spectral data and chemical composition, which are often influenced by multiple interacting factors, such as physical structure and molecular bonds. In contrast, PLS regression, a widely used linear calibration method in NIRS, was applied to model the same dataset. PLS served as a benchmark for evaluating whether the ANN model provided significant advantages over conventional linear approaches. The PLS method was applied using the same dataset and spectral pretreatment.

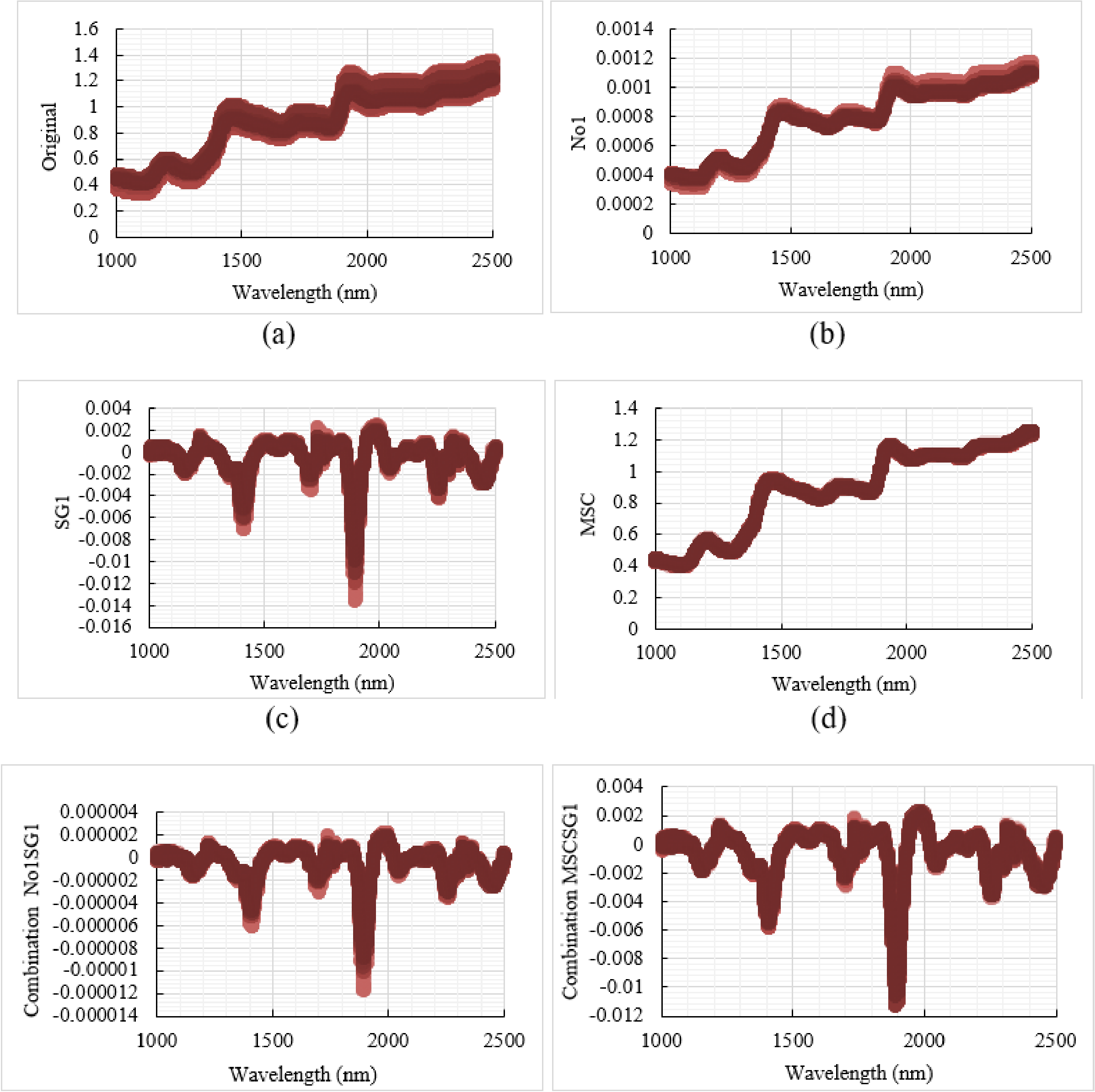
Several statistical performance parameters were used to evaluate the accuracy and reliability of the calibration and validation models: correlation coefficient (r), standard error of calibration (SEC), standard error of prediction (SEP), coefficient of variation (CV), ratio of performance to deviation (RPD), and consistency. The ideal values of these parameters are $r \approx 1$, $SEC \approx 0$, $SEP \approx 0$, $RPD \geq 2$, and $CV \approx 0$ (Nicola et al. 2007). Consistency was used as an additional parameter to assess the model stability. Consistency was calculated by comparing the SEC and SEP values, expressed as a percentage. The acceptable consistency range is 80–110% (Lengkey et al., 2013). A consistency value greater than 110% indicates a potential overfitting, whereas a value lower than 80% suggests the presence of abnormal measurement errors. Internal validation was performed to evaluate the performance and reliability of the prediction model.

3. Results and Discussion

3.1 NIRS Spectra of Robusta Coffee

The spectral data obtained from the measurements contain information about chemical components such as moisture, protein, fat, ash, and carbohydrates. Chemical content and particle size are among the factors that influence spectral shape (Aditama, 2019). According to Budiastra et al. (2018) moisture absorption occurs in the wavelength range of 1428.6-1492.5 nm, and carbohydrates in 1210-1780 nm. Protein in 1333.3–1388.9 nm, and fat in 1639.4-2325.6 nm (Aditama, 2019).

The original NIRS spectra are shown in Figure 2a. The spectra covered a wavelength range of 1000–2500 nm, with major peaks appearing at approximately 1200, 1450, 1950, and 2300 nm. These peaks are generally associated with the overtone and combination bands of the chemical bonds in coffee samples. Overall, the absorbance trend shows a gradual increase across the wavelength range, with stronger absorption observed in the higher-wavelength region.



(a) (b)

(c) (d)

(e) (f)

Figure 2. NIRS Spectra: (a) Original, (b) No1, (c) SG1, (d) MSC, (e) Combination No1SG1, and (f) Combination MSCSG1.

To maximize information extraction from NIRS spectra, transformation techniques are required to address spectral variation and overlap, achieved through pretreatment before modeling (Andasuryani et al., 2014). The No1 pre-treatment spectrum is shown in Figure 2b. This pre-treatment produces a pattern similar to the original spectrum, but with a smaller value range. This indicates that the No1 pre-treatment successfully reduced non-chemical variations, such as scattering effects, while preserving the relevant chemical information (Heil & Schmidhalter, 2021). The second pretreatment, SG1 in Figure 2c, resulted in a narrower spectral pattern compared to the original. This is because SG1 pre-treatment separates overlapping chemical components (Lengkey et al., 2013). SG1 pretreatment effectively sharpens the peaks and valleys in the NIRS spectra (Kurniasari et al., 2017). MSC pretreatment improves the spectral pattern by making the spectral bands smoother and more compact, as shown in Figure 2d. MSC pretreatment aims to reduce scattering effects to achieve the closest possible match to the standard spectrum (Ramadhan et al., 2016). Combinations of No1 pre-treatment or MSC pre-treatment with SG1 pre-treatment produced a pattern nearly identical to that of SG1 pre-treatment alone (Figures 2e and 2f). This is because the combination is predominantly influenced by SG1 pretreatment.

3.2 Chemical Content of Coffee Beans

The chemical contents of robusta coffee beans from Pagar Alam are presented in Table 1. The chemical composition of Robusta coffee beans from Pagar Alam differs from that of beans from Lampung. According to Towaha et al. (2014), the moisture content of Pagar Alam coffee ranges from 5.50%-10.30%, which is lower than that of Lampung coffee, reported at 10%-12%. Pagar Alam coffee also contains higher levels of protein (14.46–18.52%) compared to Lampung (14.64–14.94%), as well as higher fat content (4.10–6.19% compared to 1.77–2.93%) and ash content (4.74–6.16% compared to 3.73–4.19%). Meanwhile, the carbohydrate content was relatively similar, with Pagar Alam coffee ranging from 60.14-68.63% and Lampung coffee from 64.68-67.89%. These differences in chemical composition can be attributed to variations in origin, variety, and postharvest conditions (Zhang et al., 2013).

Table 1. Chemical content of Pagar Alam robusta coffee beans.

Chemical content (%)	Mean	Standard deviation	Min	Max
Moisture	7.49	1.21	5.50	10.30
Protein	17.20	0.93	14.46	18.52
Fat	5.19	0.50	4.10	6.19
Ash	5.30	0.45	4.74	6.16
Carbohydrate	64.82	1.77	60.14	68.63

3.3 Calibration and Validation NIRS for Chemical Content Prediction using PLS and PCA-ANN

The results of moisture content prediction using the PLS method are shown in Table 2, whereas the results from the PCA-ANN method are presented in Table 3. The relationship plots between the predicted moisture content (NIRS data) and the reference data (chemical data) are shown in Figure 3. The best pre-treatment for both the PLS and PCA-ANN methods was the No1SG1 pre-treatment combination. This finding aligns with that of Aditama et al. (2019), who also reported that No1SG1 was the most effective pretreatment for predicting moisture content in coffee beans. The best PLS prediction used 8 PLS factors, yielding r = 0.92, SEC = 0.47%, SEP = 0.50%, CV = 6.64%, RPD = 2.41, and consistency = 94.40%. In contrast, the best PCA-ANN prediction used 8 PC, achieving r = 0.95, SEC = 0.37%, SEP = 0.41%, CV = 5.43%, RPD = 2.95, and consistency = 90.36%. The higher r and RPD values of the PCA-ANN model indicate superior performance compared to the PLS model.

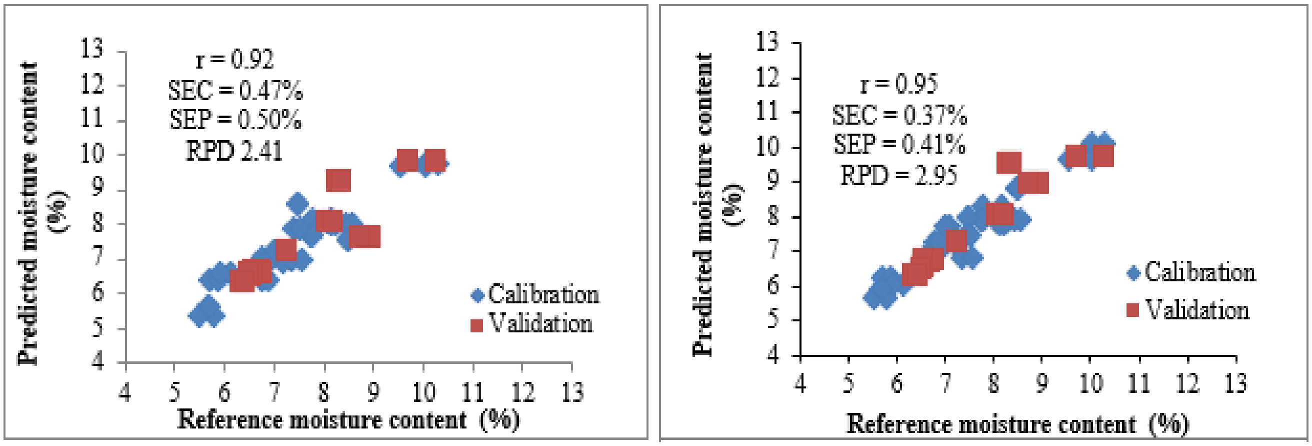
According to Nicolaï et al. (2007), RPD > 2 indicates good prediction ability, while RPD > 3 indicates very good prediction accuracy. Based on this classification, the PCA-ANN model (RPD = 2.95) falls into the good category, approaching very good, and performs better than the PLS model (RPD = 2.41), which is classified as good. According to Aditama et al. (2019), higher RPD values corresponded to better predictive models. The selection of the best model begins with identifying the highest RPD value, followed by evaluating the correlation coefficient value closest to 1 and ensuring that the consistency falls within the acceptable range of 80–110%. A good model is also characterized by a low standard error of calibration (SEC) and standard error of prediction (SEP) (Suhandy, 2010).

Table 2. Moisture content prediction results using PLS method.

Chemical Content	Pre-treatment	PLS Factor	Calibration			Validation			Consistency (%)
			r	SEC (%)	CV (%)	SEP (%)	CV (%)	RPD	
Moisture	Ori	9	0.96	0.34	4.57	1.57	20.85	0.77	21.71
		8	0.92	0.47	6.38	1.49	19.77	0.81	31.87
		7	0.86	0.62	8.36	1.19	15.83	1.01	52.37
	Nol	8	0.92	0.47	6.28	1.66	22.10	0.72	28.16
		7	0.87	0.59	7.96	1.46	19.34	0.83	40.81
		6	0.61	0.96	12.89	1.21	16.38	1.00	79.50
	SG1	8	0.92	0.48	6.38	1.49	19.78	0.81	31.99
		7	0.86	0.63	8.38	1.19	15.83	1.01	52.52
		6	0.64	0.94	12.52	1.12	14.94	1.07	83.13
	MSC	8	0.93	0.44	5.94	1.39	18.40	0.87	32.01
		7	0.88	0.58	7.74	1.38	18.31	0.87	41.93
		6	0.84	0.67	8.97	1.21	16.02	1.00	55.49
	No1SG1	**8**	**0.92**	**0.47**	**6.30**	**1.50**	**6.64**	**2.41**	**94.40**
		7	0.87	0.60	7.98	1.46	19.36	0.83	40.86
		6	0.74	0.82	10.98	1.43	18.99	0.84	57.34
	MSCSG1	8	0.93	0.45	5.98	1.39	18.42	0.87	32.22
		7	0.88	0.58	7.78	1.38	18.31	0.87	42.12
		6	0.83	0.67	8.99	1.21	16.02	1.00	55.65

Table 3. Moisture content prediction results using ANN method.

Chemical Content	Pre-treatment	PLS Factor	Calibration			Validation			Consistency (%)
			r	SEC (%)	CV (%)	SEP (%)	CV (%)	RPD	
Moisture	Ori	10	0.99	0.19	2.55	2.12	27.90	0.57	8.95
		8	0.99	0.38	5.08	1.50	19.77	0.81	25.13
		5	0.82	0.72	9.66	0.67	8.88	1.79	106.41
	Nol	10	0.99	0.19	2.59	1.78	23.61	0.68	10.88
		8	0.94	0.43	5.72	1.72	22.90	0.70	24.77
		5	0.81	0.73	9.71	1.48	19.70	0.81	48.89
	SG1	10	0.96	0.35	4.62	1.58	20.97	0.76	21.87
		8	0.98	0.23	3.06	2.12	28.18	0.57	10.75
		5	0.54	1.03	13.74	1.19	15.74	1.02	86.58
	MSC	10	0.95	0.37	4.95	1.67	22.26	0.73	22.26
		8	0.92	0.49	6.50	1.45	19.43	0.84	33.46
		5	0.62	0.95	12.71	1.25	16.73	0.98	76.00
	No1SG1	10	0.92	0.48	6.39	1.60	21.27	0.75	29.79
		8	**0.95**	**0.37**	**4.95**	**0.41**	**5.43**	**2.95**	**90.36**
		5	0.57	1.00	13.41	1.19	15.87	1.01	83.84
	MSCSG1	10	0.94	0.43	5.76	1.45	19.21	0.83	29.75
		8	0.81	0.71	9.56	1.27	16.82	0.95	56.35
		5	0.52	1.04	13.96	1.21	16.07	1.00	86.18



(a) (b)

Figure 3. Plot of the relationship between predicted moisture content and reference data: (a) PLS and (b) PCA-ANN.

Table 4 presents the results of protein prediction using the PLS method, whereas Table 5 shows the prediction results using the PCA-ANN method. A plot illustrating the relationship between the predicted protein content and the reference data is shown in Figure 4. The No1SG1 pre-treatment combination was identified as the best pre-treatment for both the PLS and PCA-ANN methods. This is consistent with the findings of Chen et al. (2013), who stated that combining pretreatment techniques can provide better effects than applying them separately.

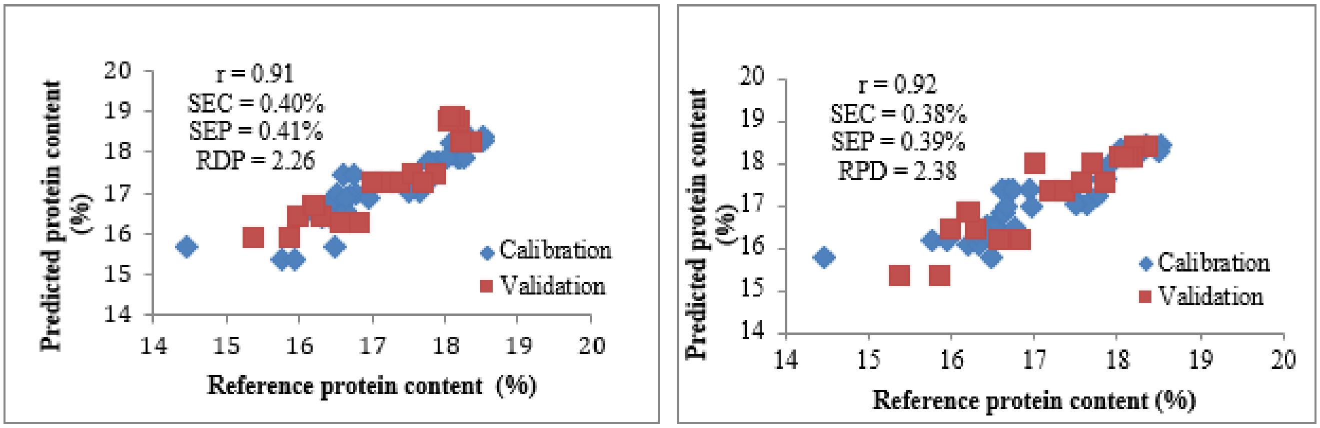
The best PLS prediction used 8 PLS factors with r = 0.91, SEC = 0.40%, SEP = 0.41%, CV = 2.40%, RPD = 2.26, and consistency = 96.30%. PCA-ANN achieved the best prediction with 8 PC yielding r = 0.92, SEC = 0.38%, SEP = 0.39%, CV = 2.28%, RPD = 2.38, and consistency = 96.78%. These results demonstrate that the PCA-ANN method outperformed the PLS method. This superiority is attributed to the strength of PCA–ANN in analyzing nonlinear relationships and identifying patterns that traditional methods such as PLS might overlook is a key advantage (Nurhadi & Hendrik, 2024).

Table 4. Protein prediction results using PLS method.

Chemical Content	Pre-treatment	PLS Factor	Calibration			Validation			Consistency (%)
			r	SEC (%)	CV (%)	SEP (%)	CV (%)	RPD	
Protein	Ori	8	0.89	0.42	2.47	0.87	5.07	1.07	48.83
		7	0.86	0.49	2.83	0.66	3.29	1.40	73.32
		6	0.78	0.59	3.43	0.67	3.88	1.77	88.75
	Nol	8	0.91	0.39	2.29	0.94	5.50	0.99	41.78
		7	0.88	0.46	2.64	0.77	4.47	1.21	59.29
		6	0.81	0.55	3.19	0.80	4.69	1.16	68.28
	SG1	8	0.89	0.43	2.47	1.17	6.81	0.80	36.47
		7	0.86	0.49	2.83	0.67	3.88	1.40	73.23
		6	0.78	0.59	3.43	0.67	3.88	1.40	88.76
	MSC	8	0.91	0.39	2.28	0.80	4.66	1.16	49.08
		7	0.89	0.44	2.53	1.06	6.17	0.88	41.11
		6	0.85	0.50	2.90	0.86	5.00	1.08	58.09
	No1SG1	**8**	**0.91**	**0.40**	**2.30**	**0.41**	**2.40**	**2.26**	**96.30**
		7	0.87	0.46	2.65	0.61	3.54	1.53	75.03
		6	0.81	0.55	3.19	0.80	4.69	1.16	68.39
	MSCSG1	8	0.91	0.39	2.29	0.80	4.69	1.16	49.33
		7	0.89	0.44	2.53	0.69	4.01	1.35	63.47
		6	0.85	0.50	2.90	0.86	5.00	1.08	58.19

Table 5. Protein prediction results using ANN method.

Chemical Content	Pre-treatment	PLS Factor	Calibration			Validation			Consistency (%)
			r	SEC (%)	CV (%)	SEP (%)	CV (%)	RPD	
Protein	Ori	10	0.87	0.46	2.70	1.29	7.49	0.72	36.12
		8	0.78	0.59	3.40	1.32	7.72	0.70	44.24
		5	0.64	0.72	4.19	0.66	3.83	1.42	109.86
	Nol	10	0.87	0.47	2.72	1.22	7.09	0.76	38.53
		8	0.79	0.58	3.36	1.04	6.06	0.89	55.67
		5	0.58	0.77	4.47	0.94	5.49	0.99	81.73
	SG1	10	0.81	0.55	3.20	1.05	6.11	0.89	52.47
		8	0.79	0.58	3.36	1.20	6.97	0.78	48.35
		5	0.70	0.67	3.92	1.02	5.93	0.91	66.23
	MSC	10	0.87	0.46	2.66	1.31	7.64	0.71	34.91
		8	0.87	0.47	2.72	1.38	8.03	0.67	33.92
		5	0.50	0.82	4.75	1.14	6.64	0.82	71.82
	No1SG1	10	0.86	0.48	2.78	0.92	5.36	1.01	52.02
		8	**0.92**	**0.38**	**2.20**	**0.39**	**2.28**	**2.38**	**96.78**
		5	0.75	0.63	3.64	0.72	4.20	1.29	86.96
	MSCSG1	10	0.87	0.46	2.68	0.80	4.66	1.16	57.74
		8	0.76	0.61	3.54	0.96	5.61	0.97	63.39
		5	0.72	0.66	3.82	0.94	5.74	0.94	66.91



(a) (b)

Figure 4. Plot of the relationship between predicted protein content and reference data: (a) PLS and (b) PCA-ANN.

The fat content prediction using the PLS method is shown in Table 6, while the results using the PCA-ANN method are presented in Table 7. A plot showing the relationship between the predicted fat content and the reference data is shown in Figure 5. The best pre-treatment for fat prediction using both the PLS and PCA-ANN methods was the No1SG1 pre-treatment. This is consistent with the findings of Aditama et al. (2019), who reported that the optimal pre-treatment for fat analysis in coffee beans was also the No1SG1 combination.

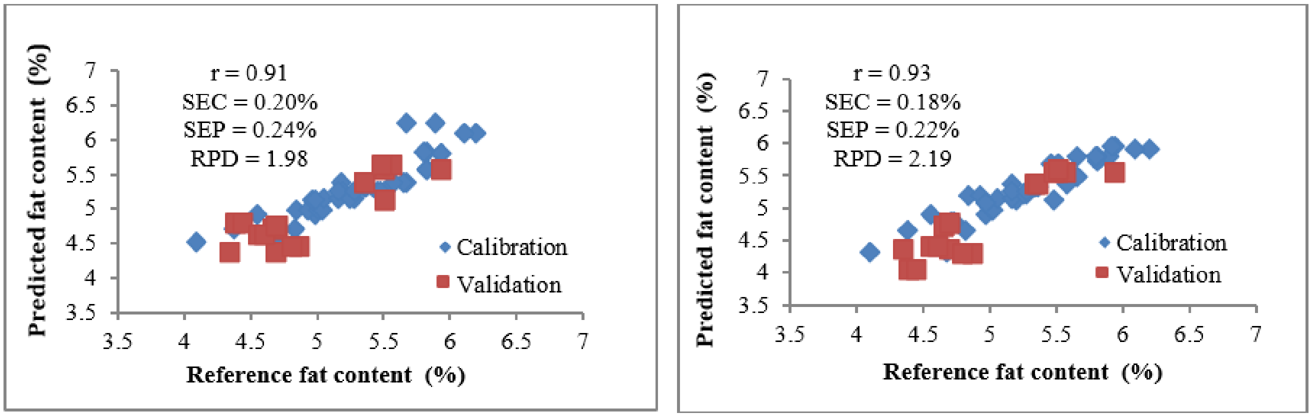
The best PLS prediction used 9 PLS factors, yielding r = 0.91, SEC = 0.20%, SEP = 0.24%, CV = 4.90%, RPD = 1.98, and consistency = 82.59%. PCA-ANN achieved the best results with 10 PC, with r = 0.93, SEC = 0.18%, SEP = 0.22%, CV = 4.43%, RPD = 2.19, and consistency = 80.08%. The PCA-ANN method demonstrated superior performance in predicting fat content in Robusta coffee beans, as indicated by the higher correlation coefficient and RPD values obtained.

Table 6. Fat prediction results using PLS method.

Chemical content	Pre-treatment	PLS Factors	Calibration			Validation			Consistency (%)
			r	SEC (%)	CV (%)	SEP (%)	CV (%)	RPD	
Fat	Ori	8	0.85	0.25	4.80	0.73	14.62	0.66	34.70
		7	0.84	0.27	5.05	0.64	12.72	0.76	41.97
		6	0.56	0.40	7.66	0.53	10.64	0.91	76.13
	No1	8	0.90	0.22	4.09	0.48	9.61	1.01	44.98
		7	0.87	0.24	4.57	0.47	9.38	1.03	51.42
		6	0.48	0.43	8.08	0.53	10.54	0.92	81.08
	SG1	8	0.85	0.25	4.82	0.73	14.63	0.66	34.84
		7	0.84	0.27	5.07	0.64	12.73	0.76	42.10
		6	0.50	0.42	8.01	0.58	11.65	0.83	72.64
	MSC	8	0.89	0.22	4.20	0.50	10.09	0.96	43.95
		7	0.87	0.24	4.59	0.67	13.48	0.72	35.95
		6	0.44	0.44	8.28	0.53	10.65	0.91	82.13
	No1SG1	**9**	**0.91**	**0.20**	**3.83**	**0.24**	**4.90**	**1.98**	**82.59**
		7	0.87	0.24	4.59	0.48	9.64	1.01	50.34
		6	0.84	0.26	5.00	0.53	10.54	0.92	50.00
	MSCSG1	8	0.89	0.22	4.23	0.53	10.65	0.91	41.97
		7	0.87	0.24	4.61	0.50	10.09	0.96	48.26
		6	0.79	0.30	5.63	0.59	11.82	0.82	50.36

Table 7. Fat prediction results using PCA-ANN method.

Chemical content	Pre-treatment	Number of PC	Calibration			Validation			Consistency (%)
			r	SEC (%)	CV (%)	SEP (%)	CV (%)	RPD	
Fat	Ori	10	0.96	0.14	2.64	1.00	19.91	0.49	14.03
		8	0.89	0.22	4.19	0.76	15.11	0.64	29.29
		5	0.80	0.30	5.62	0.67	13.37	0.73	44.44
	No1	10	0.86	0.25	4.77	0.59	11.85	0.82	42.57
		8	0.86	0.25	4.75	0.70	13.94	0.70	35.98
		5	0.87	0.24	4.59	0.71	14.29	0.68	33.90
	SG1	10	0.96	0.13	2.49	0.77	15.39	0.63	17.00
		8	0.87	0.25	4.68	0.44	8.79	1.10	56.20
		5	0.74	0.33	6.20	0.49	9.78	0.99	66.98
	MSC	10	0.87	0.24	4.60	1.06	21.24	0.46	22.88
		8	0.90	0.21	4.04	0.71	14.17	0.68	30.11
		5	0.90	0.21	3.96	0.70	14.09	0.69	29.69
	No1SG1	**10**	**0.93**	**0.18**	**3.36**	**0.22**	**4.43**	**2.19**	**80.08**
		8	0.92	0.19	3.58	0.72	14.46	0.67	26.15
		5	0.51	0.42	7.91	0.52	10.38	0.93	80.54
	MSCSG1	10	0.95	0.15	2.81	0.97	19.36	0.50	15.31
		8	0.62	0.38	7.25	0.50	10.06	0.96	76.18
		5	0.65	0.37	7.02	0.44	8.75	1.11	84.72



(a) (b)

Figure 5. Plot of the relationship between predicted fat content and reference data: (a) PLS and (b) PCA-ANN.

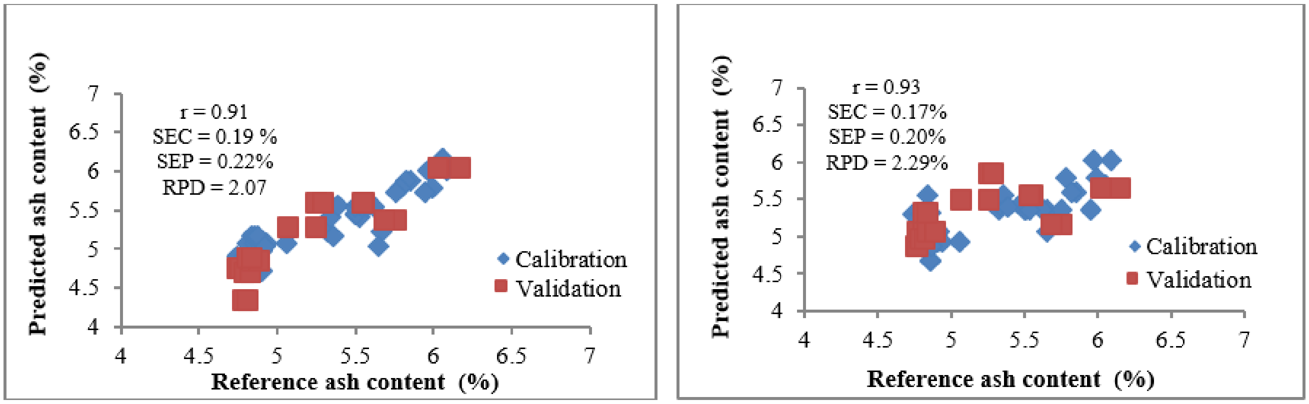
Ash content refers to the residual mineral elements left after the combustion of organic matter (Mardjan et al., 2022). Table 8 presents the prediction results of ash content using the PLS method, whereas Table 9 shows the prediction results using the PCA-ANN method. The relationship between the predicted ash content and the reference data is illustrated in Figure 6. The best PLS prediction used MSCSG1 pre-treatment with 7 PLS factors, yielding r = 0.91, SEC = 0.19%, SEP = 0.22%, CV = 4.25%, RPD = 2.07, and consistency = 84.38%. The best prediction using the PCA-ANN method was obtained with the SG1 pre-treatment and 8 PC, yielding r = 0.93, SEC = 0.17%, SEP = 0.20%, CV = 3.84%, RPD = 2.29, and consistency = 83.78%. Similar to the predictions of other chemical components in Robusta coffee beans, the ash content prediction also showed that the PCA-ANN method produced higher accuracy compared to the PLS method. The training process in ANN, which involves advanced algorithms, helps reduce prediction errors and improves the accuracy of measurement results (Nurhadi & Hendrik, 2024).

Table 8. Ash prediction results using PLS method.

Chemical content	Pre-treatment	PLS Factors	Calibration			Validation			Consistency (%)
			r	SEC (%)	CV (%)	SEP (%)	CV (%)	RPD	
Ash	Ori	8	0.90	0.20	3.66	1.09	21.11	0.42	17.89
		7	0.85	0.24	4.43	0.61	11.78	0.75	38.82
		6	0.79	0.27	5.07	0.54	10.41	0.84	50.27
	No1	8	0.91	0.18	3.44	1.15	22.13	0.40	16.08
		7	0.88	0.21	3.91	0.98	18.97	0.46	21.27
		6	0.82	0.26	4.82	0.54	10.43	0.84	47.72
	SG1	8	0.79	0.27	5.07	0.54	10.42	0.84	50.20
		7	0.85	0.24	4.43	0.61	11.78	0.75	38.87
		6	0.81	0.26	4.92	0.56	10.79	0.81	47.10
	MSC	8	0.94	0.16	2.90	1.00	19.25	0.46	15.54
		7	0.91	0.18	3.45	1.15	22.12	0.40	70.88
		6	0.82	0.26	4.78	0.64	12.30	0.71	40.16
	No1SG1	8	0.91	0.19	3.46	1.15	22.16	0.40	16.15
		7	0.88	0.21	3.92	0.98	18.97	0.46	21.33
		6	0.82	0.26	4.82	0.54	10.43	0.84	47.72
	MSCSG1	8	0.94	0.16	2.92	1.00	19.32	0.42	15.62
		7	**0.91**	**0.19**	**3.47**	**0.22**	**4.25**	**2.07**	**84.38**
		6	0.82	0.26	4.78	0.64	12.30	0.71	40.17

Table 9. Ash prediction results using PCA-ANN method.

Chemical content	Pre-treatment	Number of PC	Calibration			Validation			Consistency (%)
			r	SEC (%)	CV (%)	SEP (%)	CV (%)	RPD	
Ash	Ori	10	0.94	0.15	2.76	0.44	8.49	1.01	34.08
		8	0.94	0.15	2.83	0.45	8.73	0.98	34.04
		5	0.77	0.28	5.21	0.45	8.71	0.99	62.80
	No1	10	0.96	0.13	2.42	0.88	17.02	0.52	14.71
		8	0.94	0.15	2.83	0.45	8.73	0.98	34.04
		5	0.76	0.29	5.47	0.36	6.92	1.27	81.60
	SG1	10	0.94	0.15	2.83	0.49	9.39	0.94	31.09
		8	**0.93**	**0.17**	**3.12**	**0.20**	**3.84**	**2.29**	**83.78**
		5	0.87	0.22	4.08	0.39	7.48	1.17	56.38
	MSC	10	0.96	0.13	2.35	0.74	14.49	0.58	17.23
		8	0.94	0.14	2.67	0.56	11.07	0.76	25.65
		5	0.90	0.19	3.49	0.47	9.14	0.93	40.56
	No1SG1	10	0.95	0.14	2.65	0.38	7.24	1.21	37.85
		8	0.95	0.14	2.71	0.61	11.69	0.75	23.89
		5	0.87	0.22	4.05	0.39	7.56	1.16	55.27
	MSCSG1	10	0.94	0.15	2.79	0.34	6.54	1.34	44.09
		8	0.93	0.16	3.01	0.30	5.82	1.51	53.50
		5	0.88	0.21	3.95	0.36	6.91	1.27	59.12



(a) (b)

Figure 6. Plot of the relationship between predicted ash content and reference data: (a) PLS and (b) PCA-ANN.

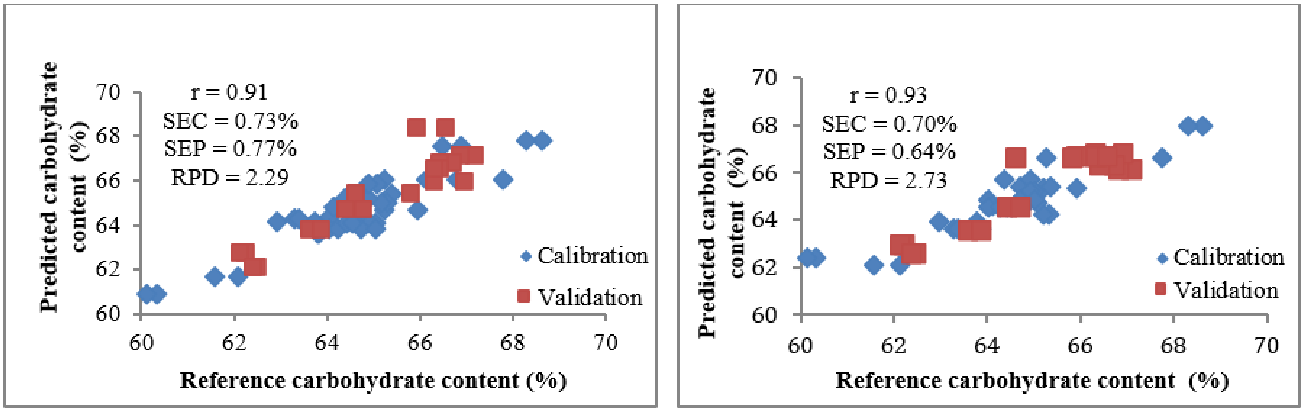
The carbohydrate content prediction results using the PLS method are shown in Table 10, whereas the prediction results using the PCA-ANN method are presented in Table 11. Figure 7 illustrates the relationship between the predicted carbohydrate content and the reference data. The best pretreatment for carbohydrate prediction using both the PLS and PCA-ANN methods was the No1SG1 pretreatment. The best prediction using PLS method was obtained with 10 PLS factors, yielding r = 0.91, SEC = 0.73%, SEP = 0.77%, CV = 1.18%, RPD = 2.29, and consistency = 94.92%. In contrast, the best prediction using PCA-ANN method was achieved with 5 PC with r = 0.93, SEC = 0.70%, SEP = 0.64%, CV = 0.99%, RPD = 2.73, and consistency = 108.85%. The correlation coefficient and RPD values from the ANN method were higher than those from the PLS method, indicating that PCA-ANN is a superior prediction model for carbohydrate content. This finding aligns with the study by Aditama et al. (2019), who concluded that the ANN method can enhance the prediction accuracy over the traditional PLS method.

Table 10. Carbohydrate prediction results using PLS method.

Chemical content	*Pre-treatment*	PLS Factors	Calibration			Validation			Consistency (%)
			r	SEC (%)	CV (%)	SEP (%)	CV (%)	RPD	
Carbohydrate	Ori	8	0.77	1.13	1.75	1.74	2.67	1.01	65.20
		7	0.60	1.41	2.19	1.97	3.02	0.89	71.83
		6	0.68	1.29	2.00	2.07	3.18	0.85	62.45
	No1	9	0.89	0.83	1.28	1.79	2.75	0.98	46.10
		8	0.86	0.92	1.42	1.39	2.14	1.26	65.81
		6	0.76	1.15	1.78	1.39	2.13	1.27	82.78
	SG1	10	0.91	0.74	1.15	1.84	2.82	0.96	40.49
		7	0.77	1.13	1.75	1.74	2.67	1.01	65.27
		6	0.68	1.29	2.00	2.07	3.18	0.85	62.54
	MSC	8	0.87	0.88	1.37	1.47	2.26	1.19	59.96
		7	0.83	1.00	1.55	1.47	2.25	1.20	68.39
		6	0.75	1.18	1.83	1.65	2.54	1.06	71.45
	No1SG1	**10**	**0.91**	**0.73**	**1.13**	**0.77**	**1.18**	**2.29**	**94.92**
		8	0.85	0.92	1.43	0.95	1.46	1.85	97.01
		6	0.76	1.15	1.78	1.47	2.26	1.19	78.21
	MSCSG1	8	0.86	0.89	1.38	1.48	2.27	1.19	60.53
		7	0.82	1.01	1.56	1.46	2.24	1.20	69.09
		6	0.81	1.05	1.62	1.33	2.03	1.33	79.21

Table 11. Carbohydrate prediction results using ANN method.

Chemical content	Pre-treatment	Number of PC	Calibration			Validation			Consistency (%)
			r	SEC (%)	CV (%)	SEP (%)	CV (%)	RPD	
Carbohydrate	Ori	10	0.89	0.83	1.28	2.40	3.69	0.74	34.46
		8	0.79	1.08	1.68	2.39	3.67	0.74	45.30
		5	0.71	1.25	1.93	1.51	2.32	1.18	82.98
	No1	10	0.87	0.88	1.35	2.45	3.75	0.72	35.82
		8	0.81	1.04	1.61	2.55	3.91	0.69	40.78
		5	0.64	1.37	2.12	1.92	2.94	0.92	71.50
	SG1	10	0.87	0.88	1.35	2.41	3.70	0.73	36.33
		8	0.82	1.03	1.59	1.87	2.87	0.94	55.01
		5	0.54	1.50	2.32	1.74	2.68	1.01	85.80
	MSC	10	0.87	0.89	1.37	1.92	2.95	0.92	46.33
		8	0.81	1.04	1.61	1.67	2.57	1.06	62.37
		5	0.71	1.25	1.42	1.99	3.06	0.89	62.88
	No1SG1	10	0.93	0.64	0.99	1.97	3.03	0.89	32.28
		8	0.87	0.90	1.39	1.88	2.88	0.94	48.06
		5	**0.93**	**0.70**	**1.08**	**0.64**	**0.99**	**2.73**	**108.85**
	MSCSG1	10	0.80	1.06	1.64	1.75	2.69	1.00	60.69
		8	0.83	1.01	1.56	1.79	2.74	0.98	56.34
		5	0.71	1.25	1.93	1.99	3.06	0.89	62.88



(a) (b)

Figure 7. Plot of the relationship between predicted Carbohydrate content and reference data: (a) PLS and (b) PCA-ANN.

4. Conclusion

The PCA-ANN method combined with spectral pre-treatment can predict the chemical content (moisture, protein, fat, ash, and carbohydrates) of robusta coffee beans more accurately than the PLS method. This is due to the ability of the PCA-ANN method to effectively handle complex nonlinear problems. PCA-ANN with No1SG1 pre-treatment yielded optimal predictions for most chemical components of robusta coffee beans, with moisture content resulting in r = 0.95, protein r = 0.92, fat r = 0.93, and carbohydrate r = 0.93, while SG1 pre-treatment was more suitable for predicting ash content with r = 0.93. These results indicate that NIRS can be used to accurately predict the chemical content of robusta coffee beans. This finding has potential applications in supporting rapid, non-destructive, and cost-effective coffee quality assessment, particularly in quality control processes within the industrial and plantation sectors of Robusta coffee in Pagar Alam.

5. References

Aditama, H. F. (2019). Development of an artificial neural network model for determining the chemical composition of Gayo Arabica coffee beans using NIRS [Bachelor's thesis, IPB University]. IPB University Repository. https://repository.ipb.ac.id/handle/123456789/98182.

Alamsyah, A., Kristanti, N., & Kristanti, F. T. (2021). Early warning model for financial distress using artificial neural network. IOP Conference Series: Materials Science and Engineering, 1098(5), 052103. https://doi.org/10.1088/1757-899x/1098/5/052103.

Andasuryani, Purwanto, Y. A., Budiastra, I. W., & Syamsu, K. (2014). Prediction of catechin content in gambir (Uncaria gambir Roxb.) using NIR Spectroscopy. Jurnal Teknologi Industri Pertanian, 24(1), 43–52.

[AOAC] Association of Official Analytical Chemists. (2006). Official method of analysis. AOAC.

Bou-Orm, N., AlRomaithi, A. A., Elrmeithi, M. (2020). Advantages of first-derivative reflectance spectroscopy in the VNIR-SWIR for the quantification of olivine and hematite. Planetary and Space Scienci, 188.

[BPS] Badan Pusat Statistik. (2024). Indonesian coffee statistics. Badan Pusat Statistik.

Budiastra, I. W., Marjan, S., Adiarifia, N., Novianty, I., & Suci, Y. T. (2024). Non-destructive prediction of oil and free fatty acid of oil palm fruitlets using near-infrared spectroscopy and hybrid calibration method. INMATEH - Agricultural Engineering, 73(2), 461–472. https://doi.org/10.35633/inmateh-73-39.

Budiastra, I. W., Sutrisno, Widyotomo, S., & Ayu, P. C. (2018). Prediction of caffeine content in java preanger coffee beans by nir spectroscopy using PLS and MLR method. IOP Conference Series: Earth and Environmental Science, 147, 012004. https://doi.org/10.1088/1755-1315/147/1/012004.

Chen, H., Song, Q., Tang, G., Feng, Q., & Lin, L. (2013). The combined optimization of savitzky-golay smoothing and multiplicative scatter correction for ft-nir pls models. ISRN Spectroscopy, 2013, 1–9. https://doi.org/10.1155/2013/642190.

Dharmawan, A., Masithoh, R. E., & Amanah, H. Z. (2023). Development of PCA-MLP model based on visible and shortwave near infrared spectroscopy for authenticating arabica coffee origins. Foods, 12(11). https://doi.org/10.3390/foods12112112.

Heil, K., & Schmidhalter, U. (2021). An evaluation of different nir-spectral pre-treatments to derive the soil parameters and of a humus-clay-rich soil. Sensors, 21(4), 1–24. https://doi.org/10.3390/s21041423.

Irawan, M., Yusmanizar, Y., & Munawar, A. A. (2022). Calibration and validation of coffee bean infrared spectral models for non-destructive testing of moisture content and chlorogenic acid. Jurnal Ilmiah Mahasiswa Pertanian, 7(4), 838–845. https://doi.org/10.17969/jimfp.v7i4.22229.

Irmeilyana, I., Ngudiantoro, N., & Rodiah, D. (2021). Correspondence analysis pada hubungan faktor-faktor yang mempengaruhi pendapatan petani kopi pagaralam. BAREKENG: Jurnal Ilmu Matematika Dan Terapan, 15(1), 179–192. https://doi.org/10.30598/barekengvol15iss1pp179-192.

Kurniasari, I., Purwanto, Y. A., Budiastra, I. W., & Ridwani, S. (2017). Prediction of tannin and total insoluble solids in persimmon fruit (Diospyros kaki L.) using NIR Spectroscopy. Jurnal Ketenikan Pertanian, 5(3), 245–252.

Lengkey, L. C. E. CH., Budiastra, I. W., Seminar, K. B., & Purwoko, B. S. (2013). Estimation model of moisture, fat, and free fatty acid content in three provenances of jatropha curcas l. seeds using near-infrared spectroscopy with the Partial Least Squares (PLS) method. Jurnal Penelitian Tanaman Industri, 19(4), 203. https://doi.org/10.21082/jlittri.v19n4.2013.203-211.

Liu, X., Fei, Y., Wang, W., Lei, S., Cheng, C., & Xing, Z. (2022). Physicochemical difference of coffee beans with different species, production areas and roasting degrees. Beverage Plant Research, 2, 1–8. https://doi.org/10.48130/BPR-2022-0007

Mardjan, S. S., Purwanto, E. H., & Pratama, Y. G. (2022). The Effect of Initial Temperature and Degree of Roasting on the Physicochemical Properties and Flavor of Solok Arabica Coffee. Jurnal Keteknikan Pertanian, 10(2), 108–122. https://doi.org/10.19028/jtep.010.2.

Masithoh, R. E., Pahlawan, M. F. R., & Wati, R. K. (2021). Non-destructive determination of SSC and pH of banana using a modular Vis/NIR spectroscopy: Comparison of Partial Least Square (PLS) and Principle Component Regression (PCR). IOP Conference Series: Earth and Environmental Science, 752(1). https://doi.org/10.1088/1755-1315/752/1/012047.

Nicolaï, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I., & Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. Postharvest Biology and Technology, 46(2), 99–118. https://doi.org/10.1016/j.postharvbio.2007.06.024.

Nurhadi & Hendrik, B. (2024). A systematic review of the role of artificial neural networks and deep learning in the diagnosis of dengue fever and typhoid. Jurnal Informatika, Manajemen Dan Komputer, 16(2), 276-288.

Olango, P. J., Lusares, R., & Yagonia-Lobarbio, C. F. (2025). A review on the potential value-added applications of extracted protein and lipids from green coffee beans. Multidisciplinary Reviews, 8(8), 1–18. https://doi.org/10.31893/multirev.2025241

Osborne, B. G., & Feran, T. H. P. (1993). Practical NIR spectroscopy with applications in food and beverage analysis. Longman Publisher.

Perera, K. D. C., Weragoda, G. K., Haputhanthri, R., & Rodrigo, S. K. (2021). Study of concentration dependent curcumin intesraction with serum biomolecules using ATR-FTIR spectroscopy combined with Principal Component Analysis (PCA) and Partial Least Square Regression (PLS-R). Vibrational Spectroscopy, 116, 103288. https://doi.org/10.1016/j.vibspec.2021.103288.

Pusat Data dan Sistem Informasi. (2023). Outlook of coffee plantation commodity. Sekertariat Jendral Kementrian Pertanian.

Ramadhan, S., Munawar, A. A., & Nurba, D. (2016). Application of NIRS and Principal Component Analysis (PCA) to detect the geographical origin of Arabica Coffee Beans (Coffea arabica). Jurnal Ilmiah Mahasiswa Pertanian, 1(1), 954–960. https://doi.org/10.17969/jimfp.v1i1.1182.

Rosita, R., Budiastra, I. W., & Sutrisno, S. (2016). Prediction of caffein content of arabica coffee bean by near infrared spectroscopy. Jurnal Keteknikan Pertanian, 04(2), 1–8. https://doi.org/10.19028/jtep.04.2.179-186

Setyawan, B., Lira, Y., & Utami, A. U. (2025). Chemical content evaluation of peaberry robusta green bean using FT-NIRS Method. 14(1), 202-2014.doi.org/10.23960/jtep-l.v14i1.202-2014.

Suhandy, D. (2010). The non-destructive method to determine soluble solids content in BW orange fruit using near infrared. Agritech, 30(1), 32–36.

Towaha, J., Aunillah, A., Heri Purwanto, E., & Handi Supriadi. (2014). The Effect of Elevation and Post-Harvest Processing on the Chemical Composition and Flavor of Lampung Robusta Coffee. J Tidp, 1(1), 57–62. https://doi.org/10.21082/jtidp.v1n1.2014.p57-62.

Yuwita, F., Ifmalinda, I., & Makky, M. (2023). Prediction of caffeine and protein of arabica coffee beans using near infrared spectroscopy (NIRS). Jurnal Teknik Pertanian Lampung, 12(4), 852. https://doi.org/10.23960/jtep-l.v12i4.852-862

Zhang, X., Li, W., Yin, B., Chen, W., Kelly, D. P., Wang, X., Zheng, K., & Du, Y. (2013). Improvement of near infrared spectroscopic (NIRS) analysis of caffeine in roasted arabica coffee by variable selection method of stability competitive adaptive reweighted sampling (SCARS). Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy, 114, 350–356. https://doi.org/10.1016/j.saa.2013.05.053.

Zhu, M., Long, Y., Chen, Y., Huang, Y., Tang, L., Gan, B., Yu, Q., & Xie, J. (2021). Fast determination of lipid and protein content in green coffee beans from different origins using NIR spectroscopy and chemometrics. Journal of Food Composition and Analysis, 102, 104055. https://doi.org/10.1016/j.jfca.2021.104055.