

RESEARCH ARTICLE OPEN ACCESS

Monitoring a Coffee Roasting Process Based on Near-Infrared and Raman Spectroscopy Coupled With Chemometrics

Leah Munyendo¹  | Katharina Schuster² | Wolfgang Armbruster² | Majharulislam Babor³ | Daniel Njoroge⁴ | Yanyan Zhang⁵ | Almut von Wrochem¹ | Alexander Schaum¹ | Bernd Hitzmann¹

¹Department of Process Analytics, University of Hohenheim, Stuttgart, Germany | ²Department of Food Chemistry and Analytical Chemistry, University of Hohenheim, Stuttgart, Germany | ³Department of Data Science in Bioeconomy, Leibniz Institute for Agricultural Engineering and Bioeconomy, Potsdam, Germany | ⁴Institute of Food Bio-resources Technology, Dedan Kimathi University of Technology, Nyeri, Kenya | ⁵Department of Flavor Chemistry, University of Hohenheim, Stuttgart, Germany

Correspondence: Leah Munyendo (leah.munyendo@uni-hohenheim.de)

Received: 1 May 2024 | **Revised:** 22 November 2024 | **Accepted:** 26 November 2024

Funding: The National Research Fund, Kenya, and the German Academic Exchange Service (DAAD), funding ID 57524989, funded this study.

Keywords: chemical changes | coffee roasting | monitoring | NIR and Raman spectroscopy

ABSTRACT

Roasting is a fundamental step in coffee processing, where complex reactions form chemical compounds related to the coffee flavor and its health-beneficial effects. These reactions occur on various time scales depending on the roasting conditions. To monitor the process and ensure reproducibility, the study proposes simple and fast techniques based on spectroscopy. This work uses analytical tools based on near-infrared (NIR) and Raman spectroscopy to monitor the coffee roasting process by predicting chemical changes in coffee beans during roasting. Green coffee beans of Robusta and Arabica species were roasted at 240°C for different roasting times. The spectra of the samples were taken using the spectrometers and modeled by the k-nearest neighbor regression (KNR), partial least squares regression (PLSR), and multiple linear regression (MLR) to predict concentrations from the spectral data sets. For NIR spectra, all the models provided satisfactory results for the prediction of chlorogenic acid, trigonelline, and DPPH radical scavenging activity with low relative root mean square error of prediction (pRMSEP < 9.649%) and high coefficient of determination ($R^2 > 0.915$). The predictions for ABTS radical scavenging activity were reasonably good. On the contrary, the models poorly predicted the caffeine and total phenolic content (TPC). Similarly, all the models based on the Raman spectra provided good prediction accuracies for monitoring the dynamics of chlorogenic acid, trigonelline, and DPPH radical scavenging activity (pRMSEP < 7.849% and $R^2 > 0.944$). The results for ABTS radical scavenging activity, caffeine, and TPC were similar to those of NIR spectra. These findings demonstrate the potential of Raman and NIR spectroscopy methods in tracking chemical changes in coffee during roasting. By doing so, it may be possible to control the quality of coffee in terms of its aroma, flavor, and roast level.

1 | Introduction

Roasting is a fundamental step in coffee processing as it induces chemical reactions, which lead to the formation of the roasted coffee's characteristic color and flavor. Green coffee

beans undergo significant chemical, physical, structural, and sensorial transformations due to multiple chemical reactions, such as the Maillard reaction, caramelization, hydrolysis, and polymerization [1]. Significant chemical changes involving compounds' degradation and formation during coffee roasting

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Author(s). *Journal of Chemometrics* published by John Wiley & Sons Ltd.

have been widely studied. These chemical changes in coffee during roasting have been experimentally determined using different off-line analytical techniques [2, 3]. The content of proteins, carbohydrates, chlorogenic acid, and trigonelline decreases considerably as the roasting degree increases, mainly associated with their involvement in the chemical reaction processes [4, 5]. On the other hand, the caffeine content is relatively stable during the roasting process [5, 6]. In addition, compounds such as melanoidins and acrylamides form through the Maillard reaction during roasting. These changes may affect the health-beneficial effects of coffee, mainly its antioxidant activity attributed to compounds such as Maillard reaction products, phenolic compounds, trigonelline, and tocopherols [7–9]. Complex reactions initiate the formation of different chemical compounds associated with the coffee flavor and its health-beneficial effects during the roasting process [10] as mentioned earlier. These reactions happen on various time scales depending on the roasting conditions. Differences in the roasting profile, the type of roaster used, and coffee bean quality, for example, size, moisture content, and chemical composition, can cause the coffee to roast differently, influencing the quality of the beverage [11]. Therefore, analytical methods to monitor the roasting process can improve the quality of the roasted coffee.

In the coffee industry, professional roasters monitor the roasting process, relying on factors such as aroma, color, and sounds emitted during a roast (cracks) [12]. This evaluation depends on the roaster's expertise and could be somewhat subjective. Furthermore, this approach may result in a lack of reproducibility and requires extensive training and experience. Because of this, spectroscopy techniques may be used as alternative tools in coffee roasting evaluation as they can be integrated into the system to monitor the process in real time and ensure accurate control and optimization of the roasting conditions, resulting in quality coffee. Existing research has shown the ability of NIR spectroscopy in combination with chemometrics to monitor acidity, color, sucrose content, and antioxidant capacity during coffee roasting [13–17]. However, few studies have assessed the potential of NIR spectroscopy to monitor the total phenolic content (TPC), antioxidant activity, chlorogenic acid, trigonelline, and caffeine content during the coffee roasting process. In addition, Raman spectroscopy has received little attention in this regard. Herdt et al. [18] utilized Raman spectroscopy to track the changes in chlorogenic acid concentration during coffee roasting. The approach has also been applied to discriminate coffee species [19, 20] and to monitor the quality of green coffee beans during storage [21].

This study aimed to develop analytical tools based on NIR and Raman spectroscopy for monitoring a coffee roasting process by predicting chemical changes (caffeine, trigonelline, chlorogenic acid, TPC, and antioxidant activity) in coffee beans. By tracking the chemical changes during roasting, it is possible to control the quality of coffee in terms of its aroma, flavor, and roast level. The roasting process can also be optimized by fine-tuning it to achieve specific chemical changes, such as enhancing desirable volatile compounds. It is also possible to create custom roast profiles based on detailed chemical data that ensure consistency in every batch. In this study, the spectra of coffee roasted at different roasting times were taken using NIR and Raman

spectrometers and modeled by the MLR, KNR, and PLSR methods to predict specific concentrations from the spectral data sets. The spectra from the two spectrometers were also combined to assess if the prediction accuracies would be improved.

2 | Materials and Methods

2.1 | Reagents and Chemicals

The caffeine standard, 2,2-diphenyl-1-picrylhydrazyl (DPPH), potassium phosphate, and 2,2'-azino-bis-(3-ethylbenzothiazolin e-6-sulfonic acid) (ABTS) were purchased from Thermo Fisher GmbH (Kandel, Germany). The standards for trolox, gallic acid, and potassium persulfate were obtained from Sigma-Aldrich Chemie GmbH (Steinheim, Germany). The following chemicals were purchased from Merck KGaA (Darmstadt, Germany): phosphoric acid, sodium carbonate, and methanol. The Folin-Ciocalteu's reagent and di-sodium hydrogen phosphate dihydrate were obtained from Carl Roth GmbH + Co. KG (Karlsruhe, Germany) and VWR International (Fontenay-sous-Bois Cedex, France), respectively. The chlorogenic acid and trigonelline standards were purchased from Cayman Chemical Company (Michigan, USA).

2.2 | Coffee Roasting, Sampling, and Grinding

Robusta and Arabica coffee beans were purchased from Buxtrade GmbH (An den Geestbergen 1, 21614 Buxtehude, Germany) and Hochland Kaffee (Hochland Kaffee Hunzelmann RS GmbH, Chemnitzer Straße 13, 70597 Stuttgart, Germany), respectively.

The beans were roasted in batches at 240°C in a Gene Cafe CBR-101 hot air coffee roaster (Gene Café, Korea) for different roast times. First, the roasting drum was preheated to the desired temperature (240°C) before the beans were poured in. Each batch contained 100 g of green coffee beans for each roast time. Coffee sampling during the roasting procedure was performed at 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, and 29 min. Each sampling meant stopping the roasting procedure at the end of each stated roasting time and collecting the whole 100 g coffee batch. This translates to 13 roastings. Four process runs were performed for each coffee species. A process run consisted of roasting green coffee beans at 240°C for different roasting times. This means for each coffee species, 52 samples were prepared (13 samples × 4 process runs). Subsequently, the roasted coffee beans were ground using an electric grinder (Melitta Calibra EU 1027-01 Mill 160 W, Germany) on a fine-grind setting. The samples were immediately stored in opaque zip-lock bags at –21°C, awaiting analysis.

2.3 | Extraction of the Samples

2.3.1 | Samples for Antioxidant Activity and Total Phenol Content Analysis

Aqueous extracts were prepared as follows: Boiled distilled water (20 mL) was added to 1 g of each sample (20:1 v/w) and extracted at 95°C for 30 min in a water bath (GFL-1083,

Burgwedel, Germany). This was followed by mixing on a vortex mixer for 1 min before centrifugation (VWR Mega Star 600R, VWR International GmbH, Darmstadt, Germany) at a speed of 4350 rpm, under a temperature of 3°C for 10 min. The resulting supernatants were then diluted accordingly before further analysis. The extracts were diluted 25-fold with a 4-mM phosphate buffer at a pH of 7.4 to determine antioxidant activity and total phenol content [22].

2.3.2 | Samples for Trigonelline, Chlorogenic Acid, and Caffeine Analysis

The extracts were prepared to analyze trigonelline, chlorogenic acid, and caffeine following the above procedure; however, 0.3 g of each sample was mixed with 40 mL of distilled water at 95°C. The undiluted extracts were filtered through a 0.45- μ m membrane (VWR, Darmstadt, Germany) before high-performance liquid chromatography (HPLC) analysis.

2.4 | Reference Measurements

2.4.1 | Antioxidant Activity Determination

2.4.1.1 | DPPH Assay. The antioxidant activity of coffee extracts by DPPH radical scavenging assay was performed using the protocol of Kwak et al. [22] with some modifications. A solution of 0.1-mM DPPH in methanol was prepared immediately before use and adjusted with methanol to an absorbance of 0.68 at 517 nm. Trolox solutions were prepared at 2, 1, 0.8, 0.6, 0.4, 0.2, 0.05, and 0.01-mM concentrations. Ten microliters of each trolox standard, diluted coffee extract sample, and phosphate buffer (blank) were pipetted into separate microtiter plate wells in duplicate. Two hundred and fifty microliters of diluted DPPH solution was added to each well using a multichannel pipette. The samples were incubated in the dark for 30 min, after which the changes in their absorbance were measured at 517 nm with a microplate reader (Infinite 200 PRO, Tecan, Männedorf, Switzerland). The results were expressed as microgram trolox equivalent (TE) per gram of coffee.

2.4.1.2 | ABTS Assay. The radical scavenging activity of coffee samples was determined by ABTS assay according to the procedure described by Xiao et al. [23] with some modifications. The ABTS radical cation was prepared by making a solution containing 7-mM ABTS and 2.45-mM potassium persulfate in 10-mL 4-mM phosphate buffer at 7.4 pH. The solution was incubated in the dark for 16 h at room temperature. The ABTS radical solution was then diluted to reach an absorbance of 0.72 at 734 nm using the phosphate buffer. Trolox solution at different concentrations (2, 1, 0.8, 0.6, 0.4, 0.2, 0.05, and 0.01 mM) was prepared and used to generate a standard curve. On a 96-well plate, 10 μ L of each diluted coffee extract sample, trolox, and buffer (blank) were pipetted into separate wells in duplicate. Subsequently, 250 μ L of diluted ABTS solution was added to each well using a multichannel pipette and incubated for 10 min. The absorbance was measured at 734 nm using a microplate reader (Infinite 200 PRO, Tecan, Männedorf, Switzerland). The ABTS

scavenging activity was expressed as microgram TE per gram of coffee.

2.4.2 | TPC Determination

The TPC of coffee extracts was determined using the Folin–Ciocalteu protocol described by Bamba et al. [24] with slight modifications. Ten microliters of gallic acid standard solutions, diluted coffee extract samples, and phosphate buffer (blank) were pipetted into separate wells in duplicate. Using a multichannel pipette, 125 μ L of 0.2-N Folin–Ciocalteu's reagent was added to each well and incubated for 10 min. Afterward, 125 μ L of 7.5% sodium carbonate was added to each well. The solutions were then incubated for 30 min, and later, the absorbance was measured at 765 nm with a microplate reader (Infinite 200 PRO, Tecan, Männedorf, Switzerland). The obtained absorbance was compared to a standard curve from gallic acid solutions (0.01–2 mM) to quantify TPC in the samples. The TPC content was expressed as microgram gallic acid equivalent (GAE) per gram of coffee.

2.4.3 | Analysis of Caffeine, Trigonelline, and Chlorogenic Acid

The content of caffeine, trigonelline, and chlorogenic acid in coffee extracts was determined using the HPLC system (UltiMate 3000, Thermo Fisher Scientific, Waltham, USA). The separation was achieved on a Nucleodur 100-3 C₈ column (150 \times 3 mm, 3 μ m; Macherey-Nagel, Dueren, Germany) at ambient temperature with UV detection at 268, 276, and 325 nm for trigonelline, caffeine, and chlorogenic acid, respectively. The mobile phases consisted of 0.01-M potassium phosphate buffer adjusted with concentrated phosphoric acid to pH 4 (Solvent A) and methanol (Solvent B). Gradient elution was performed according to Casal et al. [25]. The flow rate and the injection volume were 0.5 mL/min and 20 μ L, respectively. Duplicate measurements were obtained for each analyte.

2.5 | Spectra Acquisition

As detailed below, the off-line spectra of the roasted ground coffee samples were acquired.

2.5.1 | NIR Spectroscopy

The near-infrared (NIR) spectra of the samples were recorded using a Fourier transform NIR spectrometer (MPA, Bruker Optik, Ettlingen, Germany) equipped with a rotating sample holder. During measurement, the sample was kept in a rotation mode to ensure a collection of representative spectra. The spectra were acquired in a diffuse reflectance mode over a wavenumber range from 12,500 to 3600 cm^{-1} with a resolution of 7 cm^{-1} . Each spectrum was recorded as an average of 64 scans. The OPUS software (version 7, Bruker Optik, Germany) was used for spectra acquisition and instrumental control. In total, 52 samples for each coffee species were analyzed (13 samples \times 4 process runs). For each sample unit, five spectra were recorded,

giving a total of 520 spectra (52 samples \times 2 coffee species \times 5 replicates).

2.5.2 | Raman Spectroscopy

The Raman spectra of the samples were obtained using an Inno-Spec Raman spectrometer (Inno-Spec GmbH, Nürnberg, Germany) equipped with a process Raman probe (InPhotonics Inc., Norwood, MA, USA). All measurements were carried out at room temperature with a laser power of 500 mW and a wavelength (λ) of 785 nm. The spectra were recorded from 65 to 3290 cm^{-1} with a resolution of 1 cm^{-1} . Each spectrum was recorded as an average of 10 scans. The integration time was 1.4 s. The measurements were performed on the top of the samples where the distance between the probe (RamanProbe RPP 785/15-5 025411, InPhotonics Inc., Norwood, USA) and the sample was 3 mm. In total, 52 samples for each coffee species were analyzed (13 samples \times 4 process runs). For each sample unit, five spectra were recorded, giving a total of 520 spectra (52 samples \times 2 coffee species \times 5 replicates).

2.6 | Calibration and Validation Sets

Before further analysis, the five spectra for each sample were averaged to obtain one spectra. This means each process run had 13 samples. The NIR and Raman spectra from the four process runs, each for Arabica and Robusta coffee, were used to develop the models through a leave-one-out cross-validation approach in four iterations. In each iteration, data from three process runs were utilized to train the models, and the remaining process run was used to test their performance. The process was repeated until every process run was used as a validation set, and the average prediction accuracy was calculated. Since each process run had 13 samples after spectra averaging, 39 samples were used as a calibration set and 13 as a validation set at every iteration.

2.7 | Data Analysis

Multivariate data analysis was performed using Python (version 3.11). The collected spectra were subjected to different preprocessing methods to remove unwanted information before developing models based on MLR, KNR, and PLSR methods for predicting compounds in roasted coffee. For MLR and KNR, modeling was done using principal component analysis (PCA) extracted features referred to as principal components (PCs). The optimal number of PCs was determined by considering the total explained variance (about 95%).

2.7.1 | Preprocessing of Spectroscopic Data

Before performing statistical analysis, data preprocessing is vital in enhancing the quality of the data matrix by reducing the baseline shifts, removing scattering effects, and improving the signal-to-noise ratio [26, 27]. Different preprocessing techniques, including standard normal variate, Savitzky-Golay (SG) filter, smoothing, and their combinations, were

tested to find the best prediction accuracies of the models. Preprocessing did not improve the models' performance for NIR spectra. Therefore, the modeling was done using the raw spectra, but for visualization, the first SG derivative (2nd polynomial order, window size of 13) processed spectra are presented in Figure 1. For the Raman spectra, the accuracies were improved by removing the baseline using an iterative polynomial baseline fit method with a polynomial of order 2 and 25 iterations. After preprocessing, NIR, Raman, and combined spectra of the two spectrometers were used in modeling. A low-level data fusion strategy was used to combine the spectra of both spectrometers for evaluation. For each sample, the preprocessed data of the Raman spectrum were added to the end of the raw data of the NIR spectrum [28].

2.7.2 | Regression Models

MLR is a statistical approach that models the relationship between a dependent variable (y) and multiple independent variables (X). The association between the response and the input variables is measured by standardized regression coefficients obtained by the least squares method [29]. The approximation of y was done by a linear combination of features extracted using PCA (PCs) and the parameters.

KNR is a nonparametric method that intuitively approximates the relationship between the independent and dependent variables by averaging the observations in the same neighborhood. The algorithm searches the training data set for the k samples nearest to the sample to be classified. k is a number that indicates how many neighbors will be checked to determine the classification, and it is determined based on the distance metric [30]. This study used the Euclidean distance metric to quantify the distances. Different k values were tested, and the one that maximized the model's performance was selected. In this case, the k value was equal to three.

The PLSR method reduces the predictors to a smaller set of uncorrelated features and performs least squares regression on these features. From the spectral data, the PLS algorithm calculates new uncorrelated latent variables (LVs) that explain the maximum covariance between the spectral matrix X and the response vector y [31, 32]. The first minimum root mean square error of calibration (RMSEC) was used to determine the optimal number of LVs for regression.

2.7.3 | Model Performance Evaluation

The accuracies of the models were compared using the following quality parameters: The relative root mean square error of prediction (pRMSEP) and the coefficient of determination of prediction (R^2) were calculated as follows.

$$\text{pRMSEP} [\%] = \frac{\text{RMSEP}}{m} \times 100\% \quad (1)$$

$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}} \quad (2)$$

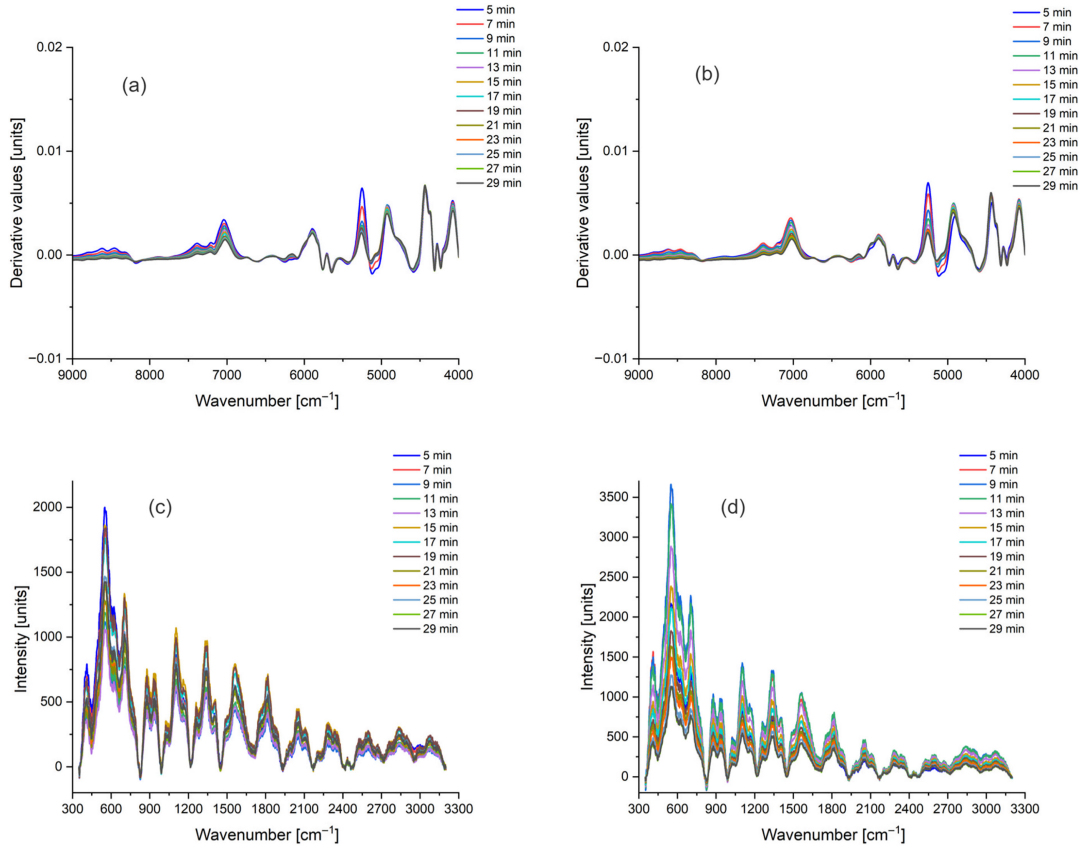


FIGURE 1 | Savitzky–Golay transformed NIR spectra for Arabica (a) and Robusta (b) coffee and baseline-corrected Raman spectra for Arabica (c) and Robusta (d) coffee roasted at different times.

TABLE 1 | The maximum, minimum, standard deviation, and mean of the compounds for the four process runs.

Compounds ($\mu\text{g/g}$)	Arabica coffee				Robusta coffee			
	Minimum	Maximum	Mean	Std Dev	Minimum	Maximum	Mean	Std Dev
Caffeine	12,250.9	15,568.5	14,536.9	533.5	13,270.5	41,193.5	26,702.5	3404.5
CGA	0.0	37,075.7	7133.0	11,137.2	191.1	47,543.4	10,477.8	14,876.3
Trigonelline	149.6	9605.8	2870.0	3133.4	118.8	7290.1	2500.2	2500.2
ABTS	46,884.7	100,114.3	67,913.7	15,230.7	60,418.6	126,675.4	88,328.0	21,113.2
DPPH	35,701.8	139,513.6	67,083.5	31,874.3	44,752.0	182,848.0	88,743.9	42,092.7
TPC	57.8	246.4	127.7	44.9	137.4	528.1	261.1	92.5

Abbreviations: ABTS, the radical scavenging activity of 2,2'-azino-bis-(3-ethylbenzothiazoline-6-sulfonic acid) expressed μg equivalent per g of coffee; CGA, chlorogenic acid; DPPH, the radical scavenging activity of 2,2-diphenyl-1-picrylhydrazyl expressed as microgram trolox equivalent per gram of coffee; Std Dev, standard deviation; TPC, total phenolic content expressed as microgram gallic acid equivalent (GAE) per gram of coffee.

$$R^2 = 1 - \frac{RSS}{TSS} \quad (3)$$

where N is the number of measurements in the test set; y_i , measured values; \hat{y}_i , predicted values; m is the difference between the maximum and minimum measured value; i , running index; TSS is the total sum of squares; and RSS is the residual sum of squares.

3 | Results and Discussion

3.1 | Spectra Overview

Figure 1a,b shows the NIR spectral profile of Arabica and Robusta coffee roasted at different times, respectively. The spectra of the samples highlight the similarity in terms of the vibrational shape and pattern. Slight differences can be

observed in the vibrational bands within regions 8200–9000, 6500–7500, and 4800–5500 cm^{-1} . These vibrational variations can be related to chemical changes during roasting, namely, the degradation or formation of compounds responsible for absorptions, which significantly affect the spectra profile. For example, the water band intensity at regions 6900 cm^{-1} (first overtone of O-H stretching) and 5200 cm^{-1} (O-H stretching and combination band of O-H deformation) decreased with an increase in roasting time as the water loss is proportional to the roast degree [33]. The second overtone of C-H stretching characterizes the vibrational band between 8200 and 9000 cm^{-1} related to carbohydrates and proteins [34]. These compounds are involved in the Maillard reaction, and thus, their decrease with an increase in roast degree could explain the lower intensities in dark-roasted samples.

Like NIR spectra, the vibrational pattern of the Raman spectra for Arabica (Figure 1c) and Robusta coffee (Figure 1d) is the same. However, slight differences in intensities among the samples can be observed along the whole wavenumber range. The bands around 2500 and 3000 cm^{-1} are attributed to the ethylenic C=C stretching and symmetric and asymmetric C-H stretching originating from the lipids [35]. The lipid content of coffee correlates positively with the roasting degree mainly due to the loss of water and degradation of carbohydrates [36]. This could explain intensity differences among the samples. The bands in the region 350–1500 cm^{-1} bands correspond to the cyclohexane ring bend and the phenyl ring stretch originating from chlorogenic acid and their isomers [37]. During

roasting, chlorogenic acid degrades to other compounds, resulting in their decrease. Therefore, the differences in intensities among the samples may indicate variations in the chlorogenic acid content. Overall, variability in the concentrations of the compounds can be seen, as indicated by high standard deviations (Table 1).

3.2 | Prediction of the Compounds Using NIR Spectra

The MLR and KNR models were developed using four PCs for all the compounds. In the case of PLSR models, the components ranged from 5 to 15 and were chosen based on the calibration error (Table 2). All the prediction models exhibited good prediction performances for chlorogenic acid, trigonelline, and DPPH radical scavenging activity with low pRMSEP (2.422%–8.226%) and high R^2 (0.941–0.994) values for Arabica coffee (Table 2). The models predicted trigonelline the best, as indicated by low pRMSEP between 2.422% and 6.625% (Table 2). The models also demonstrated high accuracies in predicting chlorogenic acid, trigonelline, and DPPH radical scavenging activity of Robusta coffee. The pRMSEP and R^2 values ranged between 4.618% and 9.649% and between 0.915 and 0.979, respectively (Table 2). Like Arabica coffee, the models predicted trigonelline the best with low pRMSEP between 4.618% and 6.637% (Table 2).

Figures 2 and 3 show the evolution of the chemical compounds of Arabica and Robusta coffee, respectively, as the roasting time

TABLE 2 | The performance of regression models on the test set for the prediction of coffee compounds using NIR spectra.

Compound ($\mu\text{g/g}$)	Model	Arabica coffee						Robusta coffee					
		PCs	pRMSEP (%)		R^2		PCs	pRMSEP (%)		R^2			
			Mean	Std Dev	Mean	Std Dev		Mean	Std Dev	Mean	Std Dev		
CGA	PLSR	5	5.510	0.975	0.963	0.019	8	7.860	3.015	0.939	0.029		
CGA	MLR	4	6.369	0.781	0.948	0.016	4	8.387	1.120	0.915	0.029		
CGA	KNR	4	3.041	1.446	0.988	0.010	4	5.042	1.867	0.974	0.021		
Trigonelline	PLSR	7	4.605	0.677	0.980	0.007	15	6.637	2.022	0.963	0.022		
Trigonelline	MLR	4	6.625	0.586	0.958	0.008	4	5.878	0.983	0.971	0.009		
Trigonelline	KNR	4	2.422	1.249	0.994	0.005	4	4.618	2.186	0.979	0.017		
ABTS	PLSR	7	11.35	1.829	0.879	0.059	7	16.29	6.034	0.775	0.128		
ABTS	MLR	4	13.21	2.184	0.801	0.093	4	14.26	5.279	0.842	0.116		
ABTS	KNR	4	10.91	1.825	0.896	0.034	4	10.76	2.695	0.902	0.064		
DPPH	PLSR	8	7.703	1.713	0.945	0.023	9	9.649	3.937	0.916	0.051		
DPPH	MLR	4	7.908	2.627	0.941	0.038	4	6.798	0.585	0.951	0.017		
DPPH	KNR	4	8.226	4.232	0.945	0.049	4	6.593	2.317	0.960	0.022		
TPC	PLSR	8	17.88	4.178	0.613	0.306	8	36.72	14.25	0.237	0.358		
TPC	MLR	4	17.90	5.812	0.752	0.105	4	23.89	11.12	0.556	0.412		
TPC	KNR	4	14.24	5.708	0.809	0.100	4	24.48	9.122	0.514	0.379		

Abbreviations: ABTS, the radical scavenging activity of 2,2'-azino-bis-(3-ethylbenzothiazoline-6-sulfonic acid) expressed microgram trolox equivalent per gram of coffee; CGA, chlorogenic acid; DPPH, the radical scavenging activity of 2,2-diphenyl-1-picrylhydrazyl expressed as microgram trolox equivalent per gram of coffee; PCs, principal components; pRMSEP, relative root mean square error of prediction; Std Dev, standard deviation; TPC, total phenolic content expressed as microgram gallic acid equivalent (GAE) per gram of coffee.

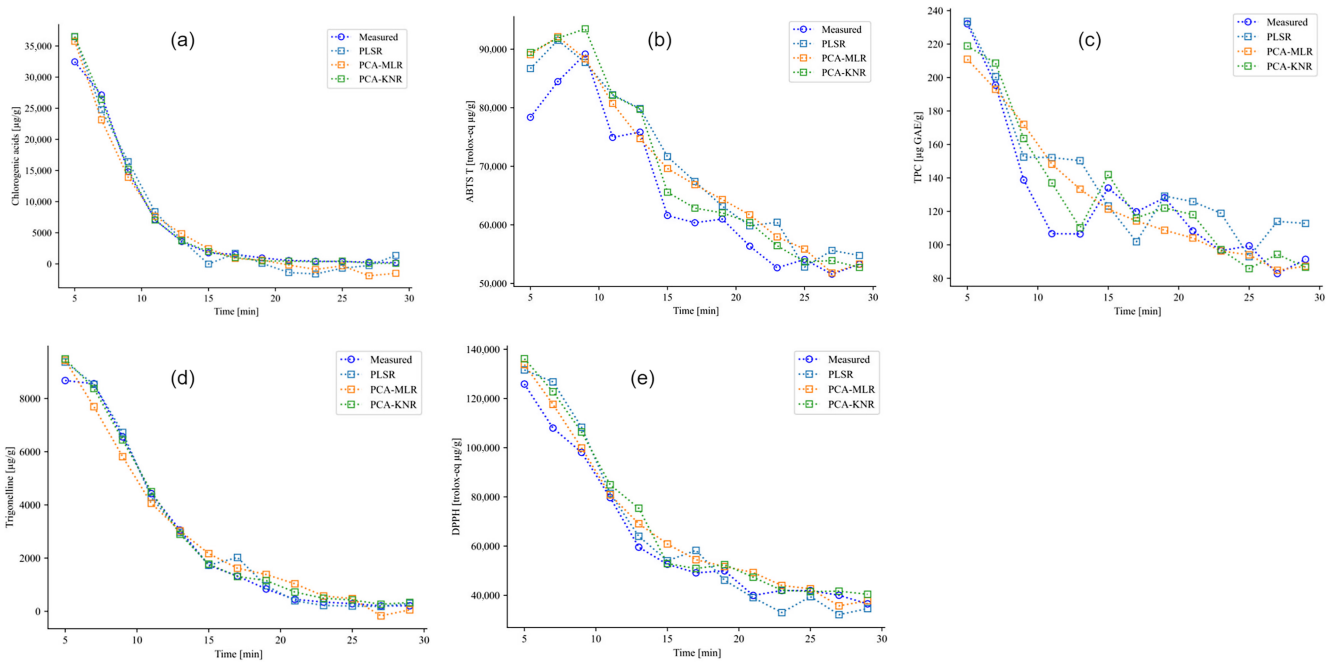


FIGURE 2 | Models prediction of the chemical compounds of Arabica coffee developed from NIR spectra (PCA followed by a regression model means that the model was calibrated and validated using PCA-extracted features).

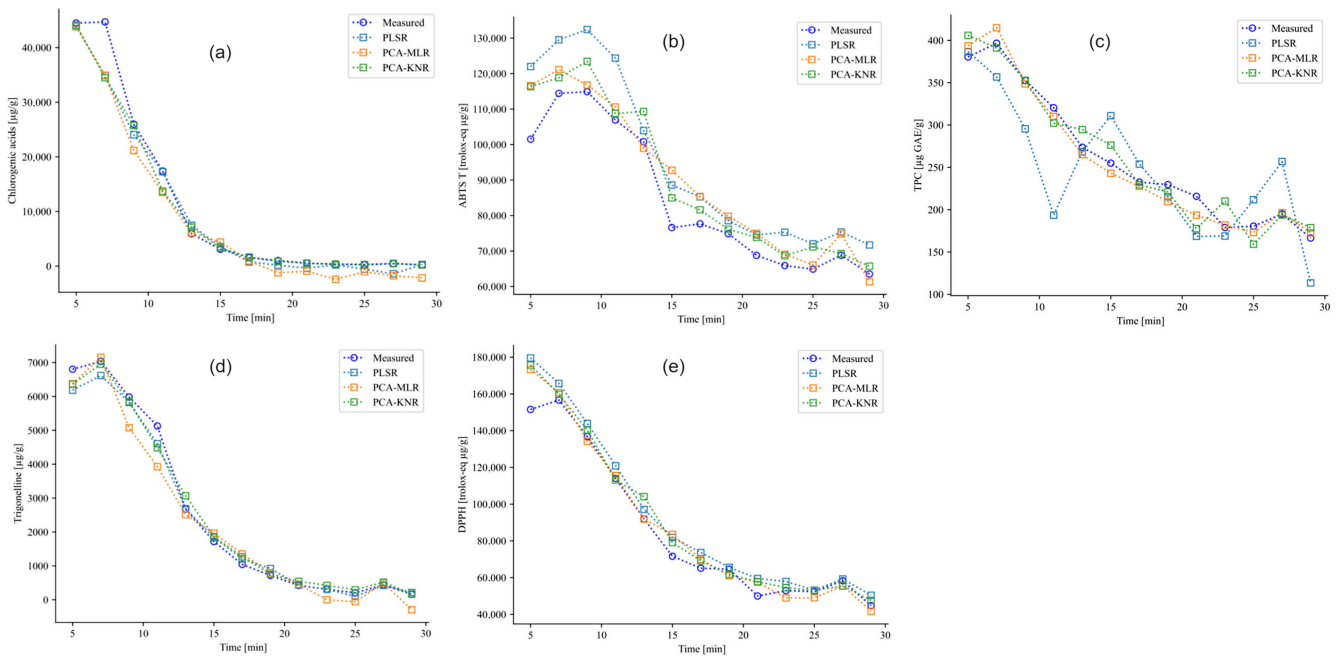


FIGURE 3 | Models prediction of the chemical compounds of Robusta coffee developed from NIR spectra (PCA followed by a regression model means that the model was calibrated and validated using PCA-extracted features).

increases predicted by different models. The models provided a good correlation between the measured and predicted content for chlorogenic acid, trigonelline, and DPPH radical scavenging activity. The excellent prediction of the compounds may be due to the models correlating their reduction over time with the complex chemical changes that occur during the roasting process. During roasting, chlorogenic acid significantly degrades into caffeic acid, lactones, and other phenol derivatives, leading to its reduction [5]. Similarly, trigonelline is broken down

to *N*-methylpyridinium, a primary contributor to the roasted beans' antioxidant activity [6]. Therefore, the loss of these compounds is directly proportional to the roast degree.

The models poorly predicted the samples' TPC and ABST radical scavenging activity, with pRMSEP above 10% for Arabica and Robusta coffee (Table 2). The high pRMSEP for the TPC may be due to its high fluctuation over roasting time, as evidenced in Figures 2c and 3c. The PLSR model had the lowest accuracy in

predicting the TPC content of Robusta coffee, with a high prediction error of 36.72% (Table 2). The ABTS radical scavenging activity correlates positively with the polyphenol content [38] and could explain the models' low prediction accuracies for the two substances. Even though the experimental conditions are not comparable, Catelani et al. [15] developed PLSR models with better prediction capabilities for TPC and ABTS radical scavenging activity ($R^2 > 0.9$). All the models poorly predicted the caffeine content with high pRMSEP (results not shown). The low accuracies of the models could be due to the irregular and minimal changes in caffeine content during roasting, as it is heat-stable [39].

Comparing the models' performance to predict the compounds in Arabica and Robusta coffee, they achieved high accuracies for Arabica coffee except for DPPH radical scavenging activity. Trigonelline content was predicted best for the two species; however, the models performed better for the Arabica coffee with lower pRMSEP (2.422%–6.625%) and higher R^2 values (0.958–0.994). For Robusta coffee, the pRMSEP and R^2 values ranged between 4.618% and 6.637% and between 0.963 and 0.979, respectively (Table 2). The model's performance was low for the TPC predictions, especially for Robusta coffee, as evidenced by pRMSEP above 20% (Table 2). Arabica and Robusta coffee species differ in their chemical composition in terms of phenolic compounds, chlorogenic acids, caffeine, trigonelline, sugars, and lipids [40, 41]. While the beans from both species undergo similar chemical transformations during roasting, the extent and nature of these changes may vary due to their distinct

initial compositions. This may cause differences in the ability of the models to identify patterns in the spectroscopic data.

3.3 | Prediction of the Compounds Using Raman Spectra

Table 3 shows the performance of regression models for the prediction of coffee compounds using Raman spectra. The models accurately predicted chlorogenic acids, trigonelline, and DPPH radical scavenging activities for Arabica coffee. The pRMSEP and R^2 values ranged between 2.286% and 7.095% and between 0.948 and 0.996, respectively (Table 3). As observed with the NIR spectra, the models predicted trigonelline content the best, with the lowest pRMSEP (2.286%–6.161%) and highest R^2 values (0.967–0.996). The models also showed robustness in predicting chlorogenic acids, trigonelline, and DPPH radical scavenging activities for Robusta coffee, achieving low pRMSEP (4.483%–7.849%) and high R^2 values (0.944–0.982). Similar to these findings, Herdt et al. [18] showed the potential of Raman spectroscopy in tracking chlorogenic acid content during coffee roasting. The models' predictive capacity for the ABTS radical scavenging activity of Arabica and Robusta coffee was fairly good. The PLSR and KNR models achieved R^2 values of 0.916 for Arabica coffee and 0.930 for Robusta coffee. On the contrary, MLR models showed low accuracies with pRMSEP above 10% and R^2 values below 0.9 (Table 3). The models performed poorly in predicting the TPC, especially for Robusta coffee, where pRMSEP of above 22% was recorded. As stated above, the

TABLE 3 | The performance of regression models on the test set for the prediction of coffee compounds using Raman spectra.

Compound ($\mu\text{g/g}$)	Model	Arabica coffee						Robusta coffee					
		PCs	pRMSEP (%)		R^2		PCs	pRMSEP (%)		R^2			
			Mean	Std Dev	Mean	Std Dev		Mean	Std Dev	Mean	Std Dev		
CGA	PLSR	5	3.027	0.501	0.991	0.002	5	4.678	1.780	0.979	0.016		
CGA	MLR	4	4.447	1.234	0.978	0.010	4	6.478	0.881	0.962	0.007		
CGA	KNR	4	3.020	0.501	0.991	0.002	4	4.661	1.784	0.979	0.016		
Trigonelline	PLSR	5	2.292	0.439	0.996	0.001	5	4.483	2.081	0.982	0.015		
Trigonelline	MLR	4	6.161	0.762	0.967	0.004	4	6.426	1.040	0.966	0.012		
Trigonelline	KNR	4	2.286	0.487	0.996	0.001	4	4.485	2.088	0.982	0.015		
ABTS	PLSR	5	10.08	1.828	0.916	0.018	5	9.384	2.924	0.930	0.033		
ABTS	MLR	4	12.14	2.443	0.872	0.030	4	11.93	2.454	0.888	0.040		
ABTS	KNR	4	10.08	1.831	0.916	0.017	4	9.371	2.927	0.930	0.033		
DPPH	PLSR	5	5.260	3.091	0.967	0.031	5	5.860	1.473	0.969	0.010		
DPPH	MLR	4	7.095	2.571	0.948	0.034	4	7.849	0.452	0.944	0.001		
DPPH	KNR	4	5.271	3.092	0.967	0.031	4	5.861	1.474	0.969	0.010		
TPC	PLSR	5	11.82	3.037	0.828	0.078	5	22.95	12.95	0.526	0.432		
TPC	MLR	4	12.71	3.486	0.792	0.098	4	23.51	12.91	0.506	0.427		
TPC	KNR	4	11.83	3.087	0.828	0.080	4	22.96	12.95	0.526	0.432		

Abbreviations: ABTS, the radical scavenging activity of 2,2'-azino-bis-(3-ethylbenzothiazoline-6-sulfonic acid) expressed microgram trolox equivalent per gram of coffee; CGA, chlorogenic acid; DPPH, the radical scavenging activity of 2,2-diphenyl-1-picrylhydrazyl expressed as microgram trolox equivalent per gram of coffee; PCs, principal components; pRMSEP, relative root mean square error of prediction; Std Dev, standard deviation; TPC, total phenolic content expressed as microgram gallic acid equivalent (GAE) per gram of coffee.

ABTS radical scavenging activity correlates positively with the polyphenol content, which could explain the models' low prediction accuracies for the two substances. However, the predictions using Raman spectra are better than those based on NIR spectra. Using Raman spectra to predict caffeine content yielded high pRMSEP and low R^2 values, like NIR spectra (results not shown).

the experimental data for chlorogenic acids, trigonelline, and DPPH radical scavenging activities. A possible explanation for this observation might be that models could identify clear patterns of these compounds in the spectroscopic data. It is important to note that the values for the PLSR and KNR models were close to each other; thus, the plots are superimposed (Figures 4 and 5). This explains the models' similar performance metrics in predicting the compounds shown in Table 3.

Figures 4 and 5 show the changes in chemical compounds of Arabica and Robusta coffee, respectively, predicted by different models. The values predicted by the models correlate well with

The models achieved higher accuracies for predicting compounds in Arabica coffee than in Robusta coffee, except for ABTS radical

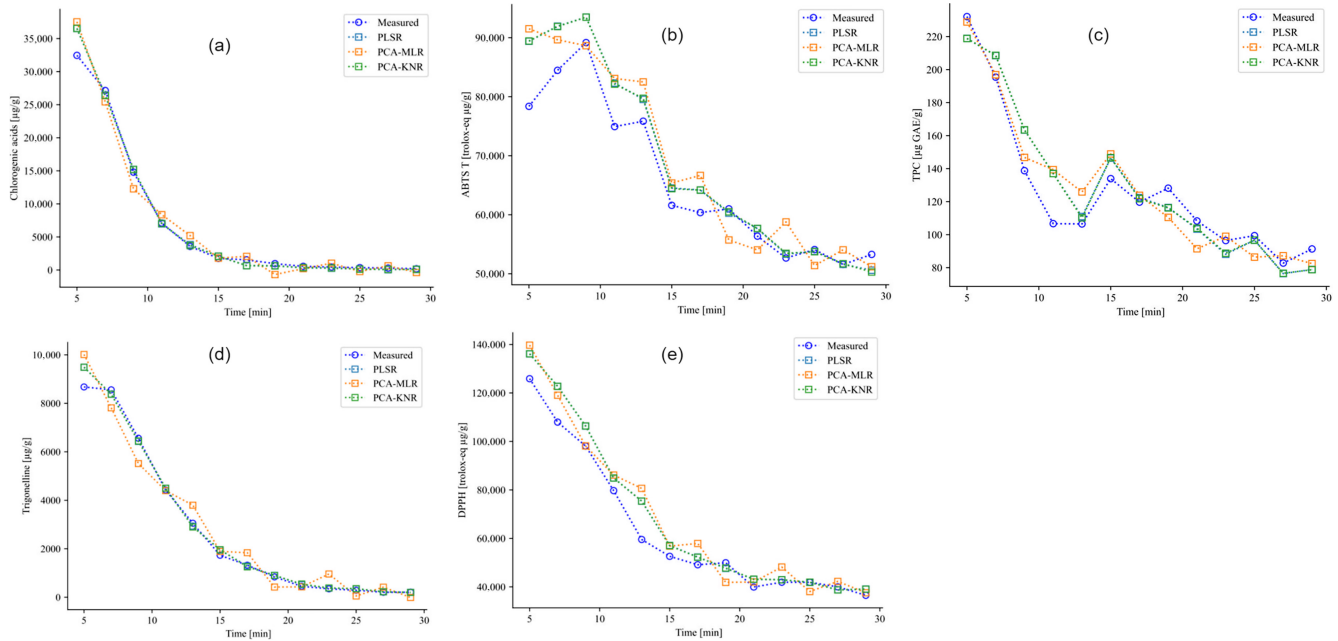


FIGURE 4 | Models prediction of the chemical compounds of Arabica coffee developed from Raman spectra (PCA followed by a regression model means that the model was calibrated and validated using PCA-extracted features).

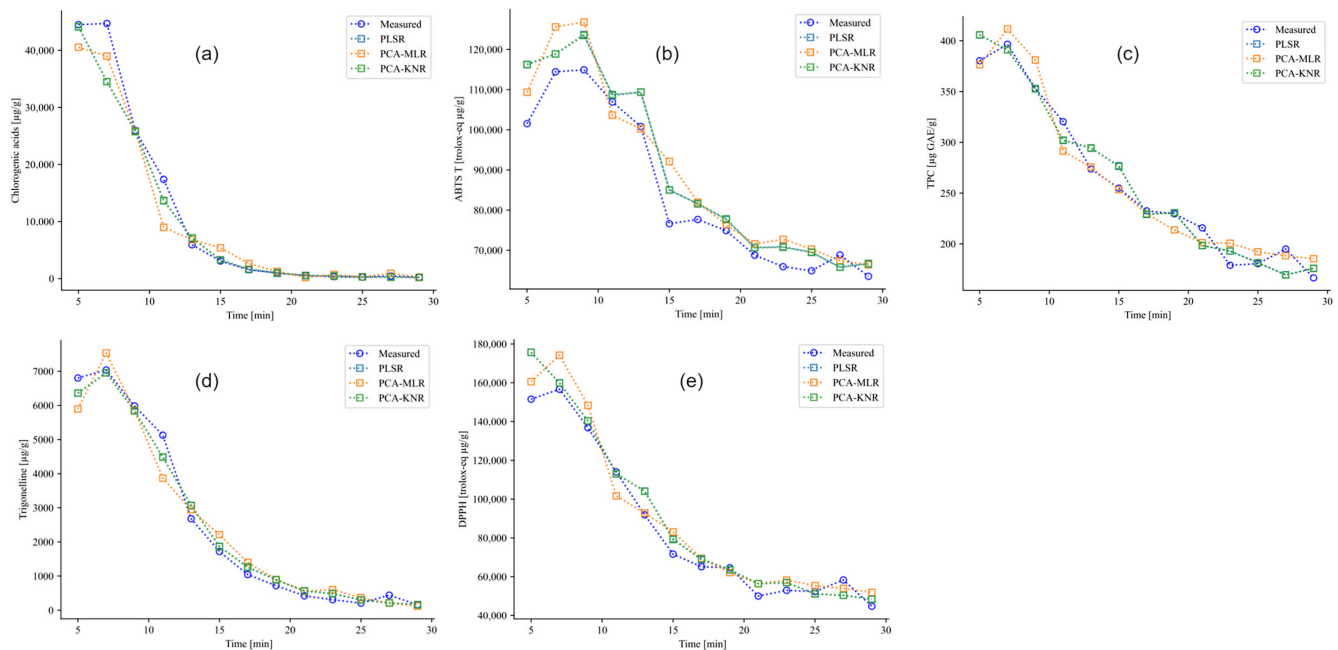


FIGURE 5 | Models prediction of the chemical compounds of Robusta coffee developed from Raman spectra (PCA followed by a regression model means that the model was calibrated and validated using PCA-extracted features).

TABLE 4 | The comparison of the model's performance for predicting coffee compounds using NIR, Raman, and combined spectra.

Compound ($\mu\text{g/g}$)	Spectrometer	Model	PCs	Arabica coffee				Robusta coffee				
				pRMSEP (%)		R^2		pRMSEP (%)		R^2		
				Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	
CGA	NIR	PLSR	5	5.510	0.975	0.963	0.019	8	7.860	3.015	0.939	0.029
		MLR	4	6.369	0.781	0.948	0.016	4	8.387	1.120	0.915	0.029
		KNR	4	3.041	1.446	0.988	0.010	4	5.042	1.867	0.974	0.021
		PLSR	5	3.027	0.501	0.991	0.002	5	4.678	1.780	0.979	0.016
	Raman	MLR	4	4.447	1.234	0.978	0.010	4	6.478	0.881	0.962	0.007
		KNR	4	3.020	0.501	0.991	0.002	4	4.661	1.784	0.979	0.016
		PLSR	5	2.662	0.457	0.993	0.002	5	4.757	1.409	0.979	0.012
		MLR	4	4.447	1.234	0.979	0.010	4	6.477	0.881	0.962	0.007
	NIR + Raman	KNR	4	3.020	0.501	0.991	0.002	4	4.661	1.784	0.979	0.016
		PLSR	7	4.605	0.677	0.980	0.007	15	6.637	2.022	0.963	0.022
		MLR	4	6.625	0.586	0.958	0.008	4	5.878	0.983	0.971	0.009
		KNR	4	2.422	1.249	0.994	0.005	4	4.618	2.186	0.979	0.017
Raman	PLSR	5	2.292	0.439	0.996	0.001	5	4.483	2.081	0.982	0.015	
	MLR	4	6.161	0.762	0.967	0.004	4	6.426	1.040	0.966	0.012	
	KNR	4	2.286	0.487	0.996	0.001	4	4.485	2.088	0.982	0.015	
	PLSR	5	2.979	0.860	0.992	0.004	5	4.392	1.894	0.983	0.013	
NIR + Raman	MLR	4	6.160	0.761	0.967	0.005	4	6.425	1.040	0.966	0.012	
	KNR	4	2.286	0.487	0.996	0.001	4	4.485	2.088	0.982	0.015	
	PLSR	5	2.979	0.860	0.992	0.004	5	4.392	1.894	0.983	0.013	
	MLR	4	6.160	0.761	0.967	0.005	4	6.425	1.040	0.966	0.012	
Trigonelline	KNR	4	2.286	0.487	0.996	0.001	4	4.485	2.088	0.982	0.015	

(Continues)

TABLE 4 | (Continued)

Compound ($\mu\text{g/g}$)	Spectrometer	Model	PCs	Arabica coffee				Robusta coffee				
				pRMSEP (%)		R^2		pRMSEP (%)		R^2		
				Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	
ABTS	NIR	PLSR	7	11.35	1.829	0.879	0.059	7	16.29	6.034	0.775	0.128
		MLR	4	13.21	2.184	0.801	0.093	4	14.26	5.279	0.842	0.116
		KNR	4	10.91	1.825	0.896	0.034	4	10.76	2.695	0.902	0.064
		PLSR	5	10.08	1.828	0.916	0.018	5	9.384	2.924	0.930	0.033
		MLR	4	12.14	2.443	0.872	0.030	4	11.93	2.454	0.888	0.040
	Raman	KNR	4	10.08	1.831	0.916	0.017	4	9.371	2.927	0.930	0.033
		PLSR	5	12.586	1.863	0.852	0.071	6	12.181	2.330	0.895	0.023
		MLR	4	12.136	2.443	0.872	0.029	4	11.930	2.454	0.888	0.040
		KNR	4	10.082	1.831	0.916	0.017	4	9.371	2.927	0.930	0.033
		PLSR	8	7.703	1.713	0.945	0.023	9	9.649	3.937	0.916	0.051
DPPH	NIR	MLR	4	7.908	2.627	0.941	0.038	4	6.798	0.585	0.951	0.017
		KNR	4	8.226	4.232	0.945	0.049	4	6.593	2.317	0.960	0.022
		PLSR	5	5.260	3.091	0.967	0.031	5	5.860	1.473	0.969	0.010
		MLR	4	7.095	2.571	0.948	0.034	4	7.849	0.452	0.944	0.001
		KNR	4	5.271	3.092	0.967	0.031	4	5.861	1.474	0.969	0.010
	NIR + Raman	PLSR	5	5.610	2.818	0.966	0.028	5	5.001	1.606	0.978	0.011
		MLR	4	7.095	2.571	0.948	0.034	4	7.848	0.451	0.944	0.010
		KNR	4	5.271	3.092	0.967	0.031	4	5.861	1.473	0.969	0.010
		PLSR	5	5.610	2.818	0.966	0.028	5	5.001	1.606	0.978	0.011
		MLR	4	7.095	2.571	0.948	0.034	4	7.848	0.451	0.944	0.010

(Continues)

TABLE 4 | (Continued)

Compound ($\mu\text{g/g}$)	Spectrometer	Model	PCs	Arabica coffee				Robusta coffee				
				pRMSEP (%)		R^2		pRMSEP (%)		R^2		
				Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	
TPC	NIR	PLSR	8	17.88	4.178	0.613	0.306	8	36.72	14.25	0.237	0.358
		MLR	4	17.90	5.812	0.752	0.105	4	23.89	11.12	0.556	0.412
	Raman	KNR	4	14.24	5.708	0.809	0.100	4	24.48	9.122	0.514	0.379
		PLSR	5	11.82	3.037	0.828	0.078	5	22.95	12.95	0.526	0.432
	NIR + Raman	MLR	4	12.71	3.486	0.792	0.098	4	23.51	12.91	0.506	0.427
		KNR	4	11.83	3.087	0.828	0.080	4	22.96	12.95	0.526	0.432
	NIR + Raman	PLSR	5	14.187	3.565	0.736	0.130	10	47.733	12.563	0.101	0.202
		MLR	4	12.708	3.486	0.792	0.098	4	23.506	12.906	0.506	0.427
		KNR	4	11.835	3.087	0.828	0.080	4	22.959	12.945	0.526	0.432

scavenging activity. Trigonelline content was predicted with the lowest pRMSEP and high R^2 values for both species. The models achieved low accuracies for TPC predictions, with Robusta coffee recording the highest pRMSEP (22.95%–23.53%) (Table 3). In the literature, Raman spectroscopy has received little attention in monitoring the chemical changes of coffee beans during roasting. Nonetheless, Abreu et al. [21] proposed a methodology based on Raman spectroscopy and multivariate control charts to identify chemical changes in green coffee beans throughout storage. A comparison of the findings from this study with the ones reported here shows the sensitivity of Raman spectroscopy in detecting chemical changes in coffee beans.

Raman spectroscopy reveals information mostly from symmetric and nonpolar molecules, whereas NIR reveals information from nonsymmetric and polar molecules [42]. These techniques complement each other and can provide additional information when used together. When the performance of the models from NIR and Raman spectra was compared, models from Raman spectra performed better in predicting all the compounds. The spectra of the two spectrometers were combined to evaluate if the prediction accuracies of the compounds would be improved. Combining the spectra did not enhance the performance of the models (Table 4). The two techniques capture complementary but fundamentally different types of information [42]. Therefore, combining them increases the diversity of the dataset, which may have led to reduced models' performance. Comparing the accuracies of the models using Raman spectra and combined spectra, Raman spectra models had better accuracies. Combined spectra models had better accuracies when the same was done for the NIR spectra.

4 | Conclusions

This investigation aimed to assess the ability of NIR and Raman spectroscopy, complemented with chemometrics, to monitor the coffee roasting process by predicting chemical changes in coffee beans during roasting. The results of this study have shown that the investigated regression models (KNR, MLR, and PLSR) based on the NIR and Raman can predict chlorogenic acid, trigonelline, and DPPH radical scavenging activity during roasting. The predictions for ABTS radical scavenging activity were reasonably good. However, the research also found low prediction accuracies for the caffeine and TPC. The KNR models showed the best prediction performances for NIR spectra, while PLSR and KNR models for Raman spectra. Overall, models from Raman spectra performed better in predicting all the compounds than those from NIR spectra. The findings reported here offer valuable insights regarding the ability of spectroscopy methods to monitor compounds in coffee during the roasting process at the industrial scale. Due to the dynamic nature of the production line in coffee processing industries, a more comprehensive or specific collection of calibration samples is necessary to develop and apply the approaches presented in this study further. Based on the results of this study, using Raman spectroscopy and PLSR or KNR methods would provide good prediction accuracies of the compounds in coffee during roasting. Future studies could assess the ability of Raman spectroscopy to predict other chemical compounds, such as proteins and carbohydrates, in coffee during roasting, as it has yet to receive much attention in this regard.

Acknowledgements

Open Access funding enabled and organized by Projekt DEAL.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Research data are not shared.

References

1. L. Poisson, I. Blank, A. Dunkel, and T. Hofmann, "The Chemistry of Roasting—Decoding Flavor Formation," in *The Craft and Science of Coffee*, 1st ed., ed. B. Folmer (Cambridge, MA: Academic Press, 2017), 273–309, <https://doi.org/10.1016/B978-0-12-803520-7.00012-8>.
2. A. A. Nuhu, "Bioactive Micronutrients in Coffee: Recent Analytical Approaches for Characterization and Quantification," *International Scholarly Research Notices* 2014 (2014): 384230, <https://doi.org/10.1155/2014/384230>.
3. A. Yashin, Y. Yashin, X. Xia, and B. Nemzer, "Chromatographic Methods for Coffee Analysis: A Review," *Journal of Food Research* 6, no. 4 (2017): 60, <https://doi.org/10.5539/jfr.v6n4p60>.
4. L. Macheiner, A. Schmidt, and H. K. Mayer, "A Novel Basis for Monitoring the Coffee Roasting Process: Isomerization Reactions of 3-Caffeoylquinic and 4-Caffeoylquinic Acids," *LWT—Food Science and Technology* 152 (2021): 112343, <https://doi.org/10.1016/j.lwt.2021.112343>.
5. S. Awwad, R. Issa, L. Alnsour, D. Albals, and I. Al-Momani, "Quantification of Caffeine and Chlorogenic Acid in Green and Roasted Coffee Samples Using HPLC-DAD and Evaluation of the Effect of Degree of Roasting on Their Levels," *Molecules* 26 (2021): 7502, <https://doi.org/10.3390/molecules26247502>.
6. A. Vignoli, C. Viegas, D. Bassoli, and M. T. Benassi, "Roasting Process Affects Differently the Bioactive Compounds and the Antioxidant Activity of Arabica and Robusta Coffees," *Food Research International* 61 (2014): 279–285.
7. N. Liang and D. Kitts, "Role of Chlorogenic Acids in Controlling Oxidative and Inflammatory Stress Conditions," *Nutrients* 8, no. 1 (2016): 16, <https://doi.org/10.3390/nu8010016>.
8. J. Folwarczna, A. Janas, M. Pylik, et al., "Effects of Trigonelline, an Alkaloid Present in Coffee, on Diabetes-Induced Disorders in the Rat Skeletal System," *Nutrients* 8, no. 3 (2016): 133, <https://doi.org/10.3390/nu8030133>.
9. A. Iriondo-DeHond, A. Rodríguez Casas, and M. D. del Castillo, "Interest of Coffee Melanoidins as Sustainable Healthier Food Ingredients," *Frontiers in Nutrition* 8 (2021): 730343, <https://doi.org/10.3389/fnut.2021.730343>.
10. M. Münchow, J. Alstrup, I. Steen, and D. Giacalone, "Roasting Conditions and Coffee Flavor: A Multi-Study Empirical Investigation," *Beverages* 6, no. 2 (2020): 29, <https://doi.org/10.3390/beverages6020029>.
11. D. G. Debona, R. F. Louvem, J. M. R. d. Luz, et al., "Heat and Mass Transfer Kinetics on the Chemical and Sensory Quality of Arabica Coffee Beans," *Agronomy* 12, no. 11 (2022): 2880, <https://doi.org/10.3390/agronomy12112880>.
12. N. Yergenson and D. E. Aston, "Monitoring Coffee Roasting Cracks and Predicting With In Situ Near-Infrared Spectroscopy," *Journal of Food Process Engineering* 43, no. 1 (2020): 13305, <https://doi.org/10.1111/jfpe.13305>.
13. E. Bertone, A. Venturello, A. Giraud, G. Pellegrino, and F. Geobaldo, "Simultaneous Determination by NIR Spectroscopy of the Roasting Degree and Arabica/Robusta Ratio in Roasted and Ground Coffee," *Food Control* 59 (2016): 683–689, <https://doi.org/10.1016/j.foodcont.2015.06.055>.
14. J. R. Santos, O. Viegas, R. N. M. J. Páscoa, I. M. P. L. V. O. Ferreira, A. O. S. S. Rangel, and J. A. Lopes, "In-Line Monitoring of the Coffee Roasting Process With Near Infrared Spectroscopy: Measurement of Sucrose and Colour," *Food Chemistry* 208 (2016): 103–110, <https://doi.org/10.1016/j.foodchem.2016.03.114>.
15. T. A. Catelani, R. N. M. J. Páscoa, J. R. Santos, et al., "A Non-Invasive Real-Time Methodology for the Quantification of Antioxidant Properties in Coffee During the Roasting Process Based on Near-Infrared Spectroscopy," *Food and Bioprocess Technology* 10, no. 4 (2017): 630–638, <https://doi.org/10.1007/s11947-016-1843-6>.
16. N. Yergenson and D. E. Aston, "Online Determination of Coffee Roast Degree Toward Controlling Acidity," *Journal of Near Infrared Spectroscopy* 28, no. 4 (2020): 175–185, <https://doi.org/10.1177/0967033520924493>.
17. F. C. Pires, R. G. F. A. Pereira, M. R. Baqueta, P. Valderrama, and R. A. Rocha, "Near-Infrared Spectroscopy and Multivariate Calibration as an Alternative to the Agtron to Predict Roasting Degrees in Coffee Beans and Ground Coffees," *Food Chemistry* 365 (2021): 130471, <https://doi.org/10.1016/j.foodchem.2021.130471>.
18. D. Herdt, T. Teumer, S. P. Keck, et al., "Quantitative Analysis of Chlorogenic Acid During Coffee Roasting via Raman Spectroscopy," *Chemosensors* 12 (2024): 106, <https://doi.org/10.3390/%0Achemosensors12060106>.
19. A. Keidel, D. Stetten, C. Rodrigues, C. Guas, and P. Hildebrandt, "Discrimination of Green Arabica and Robusta Coffee Beans by Raman Spectroscopy," *Journal of Agricultural and Food Chemistry* 58, no. 21 (2010): 11187–11192, <https://doi.org/10.1021/jf101999c>.
20. A. S. Luna, A. P. DaSilva, C. S. DaSilva, I. C. A. Lima, and J. S. deGois, "Chemometric Methods for Classification of Clonal Varieties of Green Coffee Using Raman Spectroscopy and Direct Sample Analysis," *Journal of Food Composition and Analysis* 76 (2019): 44–50, <https://doi.org/10.1016/j.jfca.2018.12.001>.
21. G. F. Abreu, F. M. Borém, L. F. C. Oliveira, M. R. Almeida, and A. P. C. Alves, "Raman Spectroscopy: A New Strategy for Monitoring the Quality of Green Coffee Beans During Storage," *Food Chemistry* 287 (2019): 241–248, <https://doi.org/10.1016/j.foodchem.2019.02.019>.
22. H. S. Kwak, S. Ji, and Y. Jeong, "The Effect of Airflow in Coffee Roasting for Antioxidant Activity and Total Polyphenol Content," *Food Control* 71 (2017): 210–216, <https://doi.org/10.1016/j.foodcont.2016.06.047>.
23. F. Xiao, T. Xu, B. Lu, and R. Liu, "Guidelines for Antioxidant Assays for Food Components," *Food Frontiers* 1, no. 1 (2020): 60–69, <https://doi.org/10.1002/fft2.10>.
24. B. S. B. Bamba, J. Shi, C. C. Tranchant, S. J. Xue, C. F. Forney, and L.-T. Lim, "Influence of Extraction Conditions on Ultrasound-Assisted Recovery of Bioactive Phenolics From Blueberry Pomace and Their Antioxidant Activity," *Molecules* 23, no. 7 (2018): 1685, <https://doi.org/10.3390/molecules23071685>.
25. S. Casal, M. B. Oliveira, and M. A. Ferreira, "Development of an HPLC/Diode-Array Detector Method for Simultaneous Determination of Trigonelline, Nicotinic Acid, and Caffeine in Coffee," *Journal of Liquid Chromatography and Related Technologies* 21, no. 20 (1998): 3187–3195, <https://doi.org/10.1080/10826079808001267>.
26. A. Rinnan, F. Berg, and S. B. Engelsen, "Review of the Most Common Pre-Processing Techniques for Near-Infrared Spectra," *Trends in Analytical Chemistry* 28, no. 10 (2009): 1201–1222, <https://doi.org/10.1016/j.trac.2009.07.007>.
27. A. Rinnan, "Pre-Processing in Vibrational Spectroscopy—When, Why and How," *Analytical Methods* 6 (2014): 7124–7129, <https://doi.org/10.1039/C3AY42270D>.

28. M. Guo, K. Wang, H. Lin, L. Wang, L. Cao, and J. Sui, "Spectral Data Fusion in Nondestructive Detection of Food Products: Strategies, Recent Applications, and Future Perspectives," *Comprehensive Reviews in Food Science and Food Safety* 23, no. 1 (2024): e13301, <https://doi.org/10.1111/1541-4337.13301>.
29. R. González-Domínguez, A. Sayago, and Á. Fernández-Recamales, "An Overview on the Application of Chemometrics Tools in Food Authenticity and Traceability," *Foods* 11 (2022): 3940.
30. D. Saha and A. Manickavasagan, "Machine Learning Techniques for Analysis of Hyperspectral Images to Determine Quality of Food Products: A Review," *Current Research in Food Science* 4 (2021): 28–44, <https://doi.org/10.1016/j.crfs.2021.01.002>.
31. T. Mehmood and B. Ahmed, "The Diversity in the Applications of Partial Least Squares: An Overview," *Journal of Chemometrics* 30 (2016): 4–17, <https://doi.org/10.1002/cem.2762>.
32. E. Bona, P. H. Março, and P. Valderrama, "Chemometrics Applied to Food Control," in *Food Control and Biosecurity* (Cambridge, MA: Academic Press, 2018), 105–133.
33. T. K. L. DeAraújo, R. O. Nobrega, D. D. S. Fernandes, M. C. U. DeAraújo, P. H. G. D. Diniz, and E. C. DaSilva, "Non-Destructive Authentication of Gourmet Ground Roasted Coffees Using NIR Spectroscopy and Digital Images," *Food Chemistry* 364 (2021): 130452.
34. A. Giraud, S. Grassi, F. Savorani, G. Gavoci, E. Casiraghi, and F. Geobaldo, "Determination of the Geographical Origin of Green Coffee Beans Using NIR Spectroscopy and Multivariate Data Analysis," *Food Control* 99 (2019): 137–145.
35. A. B. Rubayiza and M. Meurens, "Chemical Discrimination of Arabica and Robusta Coffees by Fourier Transform Raman Spectroscopy," *Journal of Agricultural and Food Chemistry* 53 (2005): 4654–4659.
36. X. Liu, Y. Fei, W. Wang, S. Lei, C. Cheng, and Z. Xing, "Physico-chemical Difference of Coffee Beans With Different Species, Production Areas and Roasting Degrees," *Beverage Plant Research* 2 (2022): 7, <https://doi.org/10.48130/bpr-2022-0007>.
37. P. J. Eravuchira, R. M. El-Abassy, S. Deshpande, et al., "Raman Spectroscopic Characterization of Different Regioisomers of Monoacyl and Diacyl Chlorogenic Acid," *Vibrational Spectroscopy* 61 (2012): 10–16, <https://doi.org/10.1016/j.vibspec.2012.02.009>.
38. A. R. Cho, K. W. Park, K. M. Kim, S. Y. Kim, and J. Han, "Influence of Roasting Conditions on the Antioxidant Characteristics of Colombian Coffee (*Coffea arabica* L.) Beans," *Journal of Food Biochemistry* 38 (2014): 271–280, <https://doi.org/10.1111/jfbc.12045>.
39. L. Song, T. Asare, M. Kang, and C. S. Lee, "Changes in Bioactive Compounds and Antioxidant Capacity of Coffee Under Different Roasting Conditions," *Korean Journal of Plant Resources* 36, no. 6 (2018): 704–713.
40. S. I. Mussatto, E. M. S. Machado, S. Martins, and J. A. Teixeira, "Production, Composition, and Application of Coffee and Its Industrial Residues," *Food and Bioprocess Technology* 4 (2011): 661–672, <https://doi.org/10.1007/S11947-011-0565-Z>.
41. N. Caporaso, M. B. Whitworth, S. Grebby, and I. D. Fisk, "Non-Destructive Analysis of Sucrose, Caffeine and Trigonelline on Single Green Coffee Beans by Hyperspectral Imaging," *Food Research International* 106 (2018): 193–203, <https://doi.org/10.1016/j.foodres.2017.12.031>.
42. G. S. Bumbrah and R. M. Sharma, "Raman Spectroscopy—Basic Principle, Instrumentation and Selected Applications for the Characterization of Drugs of Abuse," *Egyptian Journal of Forensic Sciences* 6, no. 3 (2016): 209–215, <https://doi.org/10.1016/j.ejfs.2015.06.001>.