BIO Web of Conferences **199**, 01001 (2025) https://doi.org/10.1051/bioconf/202519901001
SAGE-Grace 2025

Mathematical model for predicting coffee roast degree using spectral reflectance characteristics from the AS7265X sensor

Diang Sagita[1,2], *Sutrisno Suro* Mardjan[3*], *Slamet* Widodo[3], *Suparlan* Suparlan[2], and *Pradeka Brilyan* Purwandoko[2]

[1]Agricultural Engineering Sciences Study Program, IPB University, 16680 Bogor, Indonesia
[2]Research Centre for Appropriate Technology, National Research and Innovation Agency, 15311 Tangerang Selatan, Indonesia
[3]Department of Mechanical and Biosystem Engineering, IPB University, 16680 Bogor, Indonesia

Abstract. Some common methods for determining the coffee roast degree include colour measurement, moisture content analysis, and Agtron value assessment using a specialised spectrometer. This study develops a mathematical model to predict the roasting degree using a single reflectance characteristic from the AS7265X sensor. Spectral data from coffee samples were collected across different roasting stages, and mathematical models (i.e., linear, polynomial, exponential, logarithmic) with cross-validation techniques were built to establish relationships between individual reflectance features and key parameters: Agtron value, moisture content, and L* value. The results demonstrate that specific wavelengths strongly correlate with roasting degree indicators. The best performing models for each parameter were Agtron with reflectance at 860 nm using second order polynomial (R^2_{cv}= 0.985, $RMSE_{cv}$ = 4.671), Moisture content with reflectance at 460 nm using second order polynomial (R^2_{cv} = 0.838, $RMSE_{cv}$ = 0.864), and L* with reflectance at 810 nm using exponential (R^2_{cv} = 0.929, $RMSE_{cv}$ = 3.086). These findings highlight the potential of spectral data for non-destructive and real-time monitoring of the roasting process. The developed model offers an efficient tool for coffee quality control, particularly in industrial applications. Future research could explore machine learning techniques to enhance prediction accuracy using more robust and diverse datasets.

1 Introduction

The roasting process is a critical determinant of coffee quality, profoundly influencing the chemical composition, flavour profile, and sensory characteristics of the final product [1,2]. As coffee beans undergo thermal treatment, complex physicochemical transformations occur, including Maillard reactions, caramelization, and pyrolysis, which collectively dictate the roast degree [3]. Accurate determination of the roasting degree is essential for maintaining consistency in coffee production and ensuring optimal consumer acceptance [4,5]. It is also

[*] Corresponding author: trisno406@apps.ipb.ac.id

critical for achieving desired final quality, impacting attributes like colour, aroma, and moisture content. Traditionally, roast degree assessment has relied on subjective human evaluation, visual inspection of bean colour, or indirect measurements such as temperature-time profiling [4]. However, these methods are often inconsistent, labour-intensive, and ill-suited for real-time quality control in industrial settings. Therefore, a more reliable and objective approach is required to quantify the roasting degree accurately.

One of the most widely recognized methods for quantifying coffee roasting degree is the Agtron-based classification, which uses Agtron measuring instruments to provide a standardized numerical value [6]. This method has become a global standard, ensuring consistency in cupping tests and facilitating communication among coffee producers, roasters, and buyers [6]. By knowing the Agtron value, which ranges from 0 to 130, roast level of coffee can immediately determine quantitatively. In the Specialty Coffee Association (SCA) classification standard, there are eight coffee classifications based on Agtron value: extremely light, very light, light, medium light, medium, medium dark, dark, and very dark. For example, an Agtron value < 30 is a very dark level, while > 91 is included in the extremely light class. However, Agtron devices are expensive and not always accessible to small-scale coffee producers or researchers in resource-limited settings. This limitation has driven the search for alternative methods that are more affordable while maintaining accuracy.

Near-infrared (NIR) spectroscopy has emerged as a promising alternative for non-destructive and rapid assessment of coffee roasting degree [6]. However, studies mostly use commercially available instruments with purchase costs frequently exceeding 1000 USD [7]. Thus, among various NIR-based approaches, low-cost spectral sensors such as the AS7265X have gained attention due to their affordability, portability, and ease of integration into industrial and research applications. A study stated that the cost of making a prototype based on this sensor was only around 120 USD. The AS7265X sensor covers a broad spectral range and has been widely applied in food analysis, making it a viable candidate for coffee roasting assessment. This sensor works by transmitting light from 3 types of LEDs (UV LED, white LED, and IR LED) to the material/product, then its reflectance is captured by a photodiode sensor consisting of 18 wavelength channels ranging from 410 nm to 940 nm. The 410–940 nm spectral range (visible and shortwave NIR) according to the AS7265X sensor specifications has the potential to be used as a basis for estimating roasting levels, represented by parameters such as moisture content, lightness (L*) value from the CIE L*a*b* colour space, and the Agtron index. Studies have shown that the L* value of coffee decreases as roasting time increases [1,8], which is closely related to changes in the visible light spectrum. In addition, water content and Agtron values may be estimated through vibrations in both the visible and near-infrared (NIR) regions, as chemical changes in coffee during roasting are reflected in its physical appearance, particularly colour. This aligns with research indicating that wavelengths around 700 nm and 900 nm correspond to the third overtone vibrations of O–H and C–H bonds, respectively [9].

To maximize the potential of AS7265X in coffee roasting classification, there is a need for a simple yet effective mathematical model that can predict Agtron values and other key roasting parameters based on individual reflectance data. By developing such a model, the roasting degree can be estimated in a cost-effective manner, reducing dependence on expensive Agtron instruments. Therefore, this research aims to develop a mathematical model to predict coffee roast degree (Agtron, L*, and moisture content) using spectral characteristics from the AS7265X sensor and provide a low-cost and non-destructive alternative for real-time monitoring of coffee roasting processes.

2 Materials and methods

2.1 Sample preparation

The coffee samples used were separate types of Robusta and Arabica coffee. Robusta coffee came from Dampit (East Java, Indonesia) and Temanggung (Central Java, Indonesia), and Arabica came from Gayo (Aceh, Indonesia) and Flores (East Nusa Tenggara, Indonesia). All coffee beans obtained through certified distributors comply with the Indonesian National Standard for Grade 1 quality, namely coffee beans free from defects (no insects, fungi, or rot), a maximum water content of 12.5%, and impurities not exceeding 0.5% of the total mass. Before the roasting experiment, the coffee beans were stored in tightly closed containers at a controlled temperature between 25°C and 30°C.

Roasting was carried out using a drum-type coffee roaster with a capacity of 500 grams per batch. Roasting began with a preheating process so that the temperature of the roasting room reached 190°C. After this temperature was reached, the coffee sample was fed through the hopper. Roasting was performed at several roasting time variations to obtain coffee at various roasting levels, i.e., 3, 5, 7, 8, 9, 9, 9.5, 10., 10.5, and 11 minutes. Coffee parameter data were taken in three repetitions.

2.2 Analysis of Agtron, moisture content, and L* value

The analysis of coffee sample includes Agtron value, moisture content, and lightness value represented by L* which is colorimetric indicator from the CIE L*a*b* colour space. Agtron parameters was measured using a roast coffee analyser CM-100 type (Lighttells Corp. Ltd., Taiwan) which represents the roast level quantitatively with qualifications between extremely dark (<30) to extremely light (>91). Moisture content was measured using the method of the Association of Official Analytical Chemists (AOAC), which involves gravimetric drying to quantify moisture levels. L* parameter was measured via the CIE Lab* colour space system using a Konica Minolta CR-400 Chromameter (Konica Minolta Sensing, Inc., Osaka, Japan).

2.3 Spectral data acquisition

To measure spectral data from the coffee samples in the form of light reflection properties across visible and near-infrared wavelengths (410–940 nm), a specialised system was developed using an AS7275X multispectral detector [4]. This configuration gathered reflectance information via 18 unique wavelength bands, allowing a detailed examination of coffee beans' light-interaction features. The AS7265X sensor, a compact and budget-friendly tool for multi-wavelength analysis, covers 18 spectral bands within the 410–940 nm spectrum. Its affordability and small size make it advantageous for projects with constrained resources. The device integrates effortlessly with common microcontroller platforms (e.g., Arduino, ESP) using standardized communication protocols, streamlining its incorporation into experimental setups. Functionally, the system uses a tri-phase lighting approach: ultraviolet (UV), visible, and infrared (IR) LEDs activate in sequence to scan samples (Fig. 1). Six narrow bands (410, 435, 460, 485, 510, 535 nm) detected the UV LED (405 nm) reflectance, six intermediate wavelengths (560, 585, 610, 645, 680, 705 nm) captured the visible LED (5700 k) reflectance, and six longer wavelengths (730, 760, 810, 860, 900, 940 nm) detected the IR LED (875 nm) reflectance.

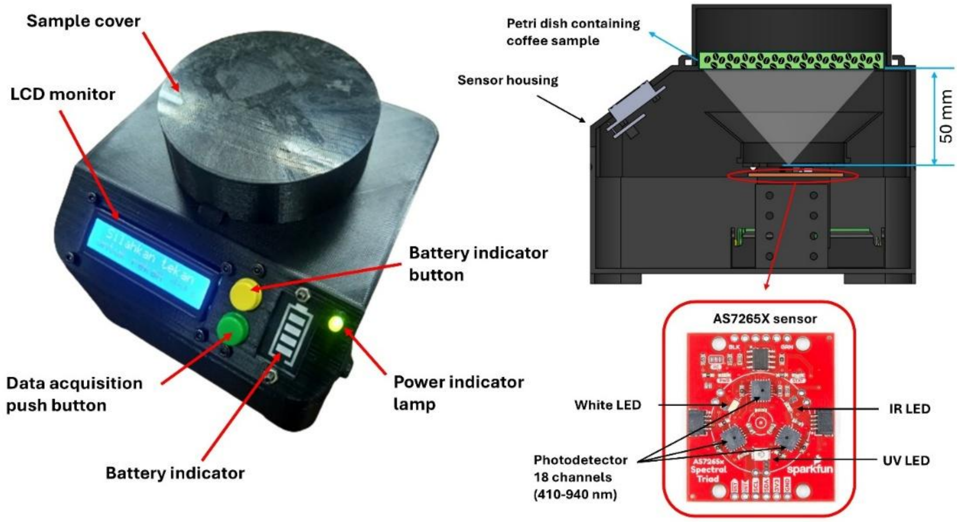

Fig. 1. Schematic principle of coffee bean spectral data acquisition using AS7265X (Adapted from [4]).

2.4 Mathematical modelling

This study aims to develop a mathematical model for predicting coffee roasting indicators, including the novel application of the Agtron scale, which is a previously underexplored variable in this context, alongside the established parameters L* (lightness) and moisture content. The modelling framework employs a single-variable regression approach to identify optimal wavelengths within spectral data for accurate estimation of these indicators. Four types of mathematical models were employed in this study: linear, second-order polynomial, exponential, and logarithmic, as presented in Eq. 1 to Eq. 4, where Y is the estimated parameter, X is the spectral data value at a certain wavelength, a is a constant, $b1$ *and* $b2$ are coefficient. The model evaluation metrics used as criteria are R^2 and $RMSE$ which are calculated using Eq. 5 and Eq. 6 respectively. Higher R^2 values, along with lower RMSE, indicate better model performance [10]. Modelling was performed using a cross-validation technique with 5K-folds as presented in Fig. 2 combined with reflectance parameter optimization using GridSearchCV. All computational work was performed using Python 3.9.13 in a Jupyter Notebook environment (v6.5.4). By systematically linking spectral patterns to roasting metrics, this approach provides a data-driven foundation for optimizing roast-level assessment in coffee processing.

$$Y = a+b1X \tag{1}$$

$$Y = a+b1X+b2X^2 \tag{2}$$

$$Y = EXP(a+b1X) \tag{3}$$

$$Y = a+b\ ln(X) \tag{4}$$

$$R^2_{cv} = 1-\frac{\sum_{i=1}^{N}(y_{e,i}-y_{p,i})^2}{\sum_{i=1}^{N}(y_{e,i}-\overline{y_i})^2} \tag{5}$$

$$RMSE_{cv} = \left[\frac{1}{N}\sum_{i=1}^{N}\left(y_{e,i}-y_{p,i}\right)^2\right]^{\frac{1}{2}} \tag{6}$$

 https://doi.org/10.1051/bioconf/202519901001

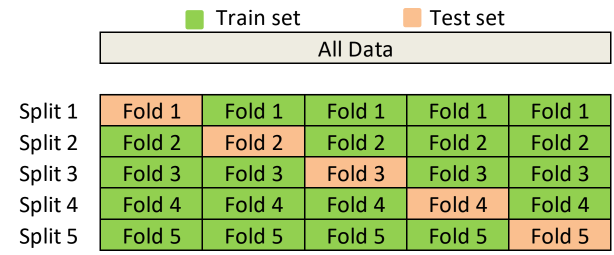

Fig. 2. Model evaluation using 5K-fold cross-validation.

3 Results and discussion

3.1 Effect of roasting on Agtron, moisture content, and L* value

Fig. 3 presents three sub-graphics illustrating the effect of roasting time (x-axis in all cases) on three different quality attributes of coffee: Agtron value, moisture content, and L* value. The Agtron value, which indicates roast degree (higher values = lighter roast), shows a strong nonlinear decrease as roasting time increase. The parabolic model (concave downward) suggests a rapid decline in Agtron value after a certain roasting threshold, consistent with the browning and Maillard reactions accelerating as beans are roasted longer. The high R^2 (0.9403) indicates excellent model fit, suggesting that roasting time strongly explains the variation in Agtron value. This is critical for roast control, as Agtron is a standard metric for roast classification (light, medium, dark).

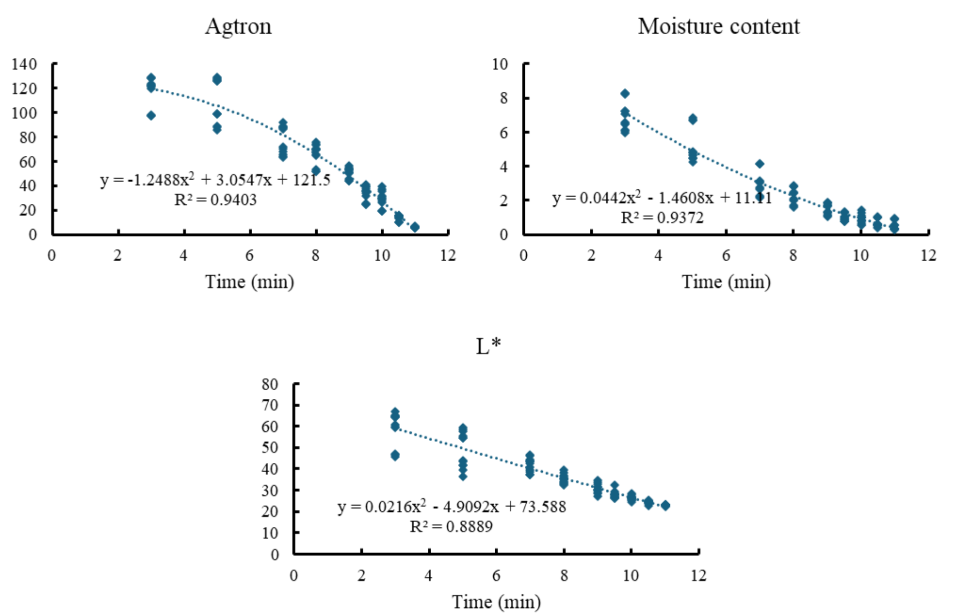

Fig. 3. Changes in Agtron, moisture content, and L* value during roasting.

Moisture content decreases with increasing roasting time, which is expected as water evaporates during thermal treatment. Although the curve is slightly nonlinear, the dominant trend is a nearly linear decrease, especially in the early to mid-stages of roasting. The slight upward curvature in the polynomial could indicate a diminishing rate of moisture loss at longer roast times, possibly due to bound water or thermal equilibrium effects. The high R^2 again confirms a strong relationship, suggesting roasting time is a reliable predictor of residual moisture—a key parameter for shelf-life and quality preservation. Previous published paper reported that coffee roasting produces significant changes in the physical properties and chemical content of coffee beans [1,11,12]. The transformation from green bean coffee to roasted coffee causes physical changes such as moisture content, shape, density, structure and colour [12].

L* value decreases with roasting time, aligning with visual darkening due to pyrolysis and caramelization. This phenomenon is in line with several literature related to the effects of roasting on coffee characteristics. It was reported that roasting darkens the colour of coffee beans, with lightness (L*) values decreasing as roasting temperature and time increase [13,14].The trend shows a nonlinear reduction, possibly indicating a faster drop in lightness at earlier roasting stages, followed by a slower decline as beans approach very dark roast levels. The R^2 value of 0.889 is slightly lower than the other models, but still indicates a strong relationship, confirming L* is sensitive to roast progression and could be used for non-invasive roast monitoring.

3.2 Mathematical model for Agtron, moisture content and L* value

Table 1 presents the model evaluation results for predicting Agtron, moisture content, and L* values based on individual spectral reflectance features (R) from the AS7265X sensor across 18 wavelengths (410–940 nm). Agtron values relate to internal chemical changes (e.g., Maillard reaction products, caramelization), which are best captured in the NIR and deep-red regions due to strong absorption by organic compounds. These regions penetrate deeper into the bean, providing information beyond surface colour. It was found that the reflectance at a wavelength of 860 nm gave the best value (R^2_{cv} = 0.985, $RMSE_{cv}$ = 4.671) with the best mathematical model obtained presented in Eq. 7. In addition, linear model also showed comparable result with R^2_{cv} 0.984. Several wavelengths with linear and polynomial model also showed R^2_{cv} values > 0.95 such as 680, 810, 900 and 940 nm.

Moisture prediction shows slightly lower R^2 values overall, with mid-visible to NIR wavelengths performing best. The top results are 460 nm using polynomial (R^2_{cv} = 0.838, $RMSE_{cv}$ = 0.864), 510 nm using logarithmic (R^2_{cv} = 0.832, $RMSE_{cv}$ = 0.870), and 510 nm using linear (R^2_{cv} = 0.831, $RMSE_{cv}$ = 0.878). Water absorbs weakly in visible wavelengths but may influence reflectance due to changes in bean structure and light scattering. Thus, the wavelength of 460 nm gives the highest value with the mathematical model presented in Eq. 8.

L* value, representing surface lightness, shows high accuracy in the red and NIR region. The best performance is 810 nm using exponential (R^2_{cv} = 0.929, $RMSE_{cv}$ = 3.086), followed by 810 nm using polynomial (R^2_{cv} = 0.928, $RMSE_{cv}$ = 3.068), 705 nm using linear and 560 nm using logarithmic. L* is a surface colour metric, and is therefore more influenced by visible reflectance, especially in the red/orange-yellow regions that correlate well with browning intensity. The best mathematical model for this parameter is presented in Eq. 9.

$$A = -59.4376+0.6825X+0.000741X^2 \quad (7)$$
$$Mc = -13.5112+1.278X+0.0196X^2 \quad (8)$$
$$L^* = EXP(2.2784+0.0246X) \quad (9)$$

The results of fitting the best model for the three parameters and the distribution of the data samples are presented in Fig. 4. It can be seen that of the three parameters, the best modelling results are for the Agtron parameter, where the data appears close to the best fit line, while the water content and L* parameters tend to be somewhat scattered. This is why the R^2 values for water content and L* are not higher than the Agtron value. From these results, it was also revealed that the Agtron value is the most suitable to be a roasting parameter because this index has become the standard for classifying roast levels into 8 classes ranging from extremely light, light, medium, dark and even very dark. For example, an Agtron value > 91 is an extremely light roast level, a value of 81-90 is very light, a value of 51-60 is medium level and so on. Furthermore, the model formed with only one spectral variable, viz. 860 nm, has been proven to be usable because the RMSE value is below 5 (Agtron scale), which is still within the range of values in a roasting level class.

Table 1. Results of mathematical model evaluation metrics after optimization using GridSearchCV on each type of model and wavelength feature.

Target	Model	Feature	R^2_{cv}	$RMSE_{cv}$
Agtron	Linear	R (860 nm)	0.984	4.754
	Polynomial	**R (860 nm)**	**0.985**	**4.671**
	Exponential	R (940 nm)	0.769	19.274
	Logarithmic	R (645 nm)	0.974	6.076
Moisture content	Linear	R (510 nm)	0.831	0.878
	Polynomial	**R (460 nm)**	**0.838**	**0.864**
	Exponential	R (535 nm)	0.751	1.081
	Logarithmic	R (510 nm)	0.832	0.870
L* value	Linear	R (705 nm)	0.915	3.492
	Polynomial	R (810 nm)	0.928	3.068
	Exponential	**R (810 nm)**	**0.929**	**3.086**
	Logarithmic	R (560 nm)	0.904	3.520

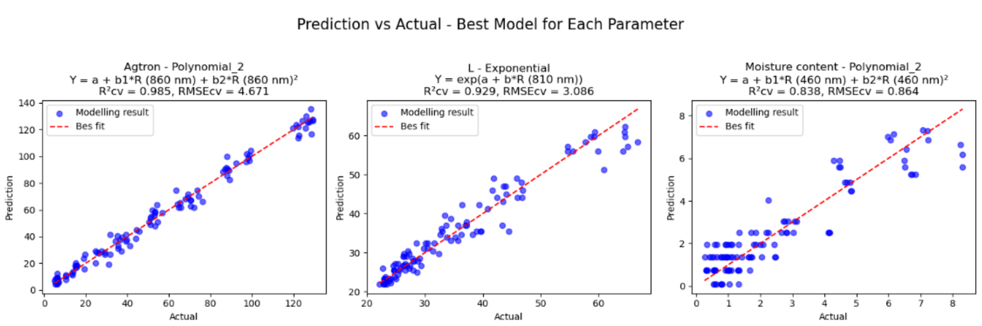

Fig. 4. The best fitting results of the mathematical model for Agtron, L* value, and moisture content.

The findings of this study highlight the significance of accurately estimating roast level as an indicator of coffee quality, as literature shows that roast level directly influences flavour [14]. Different coffee types tend to reach their optimal flavour at varying roast levels (as measured by the Agtron value), and individual preferences for roast level can also vary widely [4]. Therefore, the mathematical model developed in this study offers practical value, as it can be integrated into a low-cost prototype roast level measuring device using the AS7265X sensor, providing an affordable alternative to the high-cost Agtron instrument.

4 Conclusion

This study demonstrated the effectiveness of mathematical modelling and spectral analysis using the AS7265X multispectral sensor to predict key quality parameters of roasted coffee: Agtron value, moisture content, and L*. Polynomial regression models showed strong relationships between roasting time and each parameter, with R^2 values of 0.9403 (Agtron), 0.9372 (moisture), and 0.8889 (L*), confirming the nonlinear dynamics of thermal transformation during roasting. Further evaluation of individual spectral features revealed that specific wavelengths within the AS7265X sensor provide high predictive performance. For Agtron, the best results were achieved in the near-infrared (NIR) region, especially at 860 nm (R^2_{cv}= 0.985, $RMSE_{cv}$ = 4.671), indicating strong sensitivity to internal chemical changes such as Maillard reaction products. Moisture content showed moderate predictability, with optimal results around 460 nm (R^2_{cv} = 0.838, $RMSE_{cv}$ = 0.864), likely due to the combined effects of water loss and matrix transformation affecting light scattering. L* value, representing surface lightness, was best predicted using wavelengths in the shortwave NIR region (R^2_{cv} = 0.929, $RMSE_{cv}$ = 3.086). These findings validate the potential of low-cost spectral sensors and simple modelling approaches for rapid, non-destructive evaluation of coffee roasting quality. The distinct optimal wavelengths for each parameter also suggest opportunities for multispectral feature fusion or optimization in future model development. Overall, this approach offers a promising tool for enhancing quality control and process automation in coffee roasting industries. Further research is recommended to develop a more accurate and generalizable model. This can be achieved by applying machine learning techniques to a robust dataset that includes diverse coffee varieties.

Acknowledgements. This work was financially supported by Home Program in Research Organization for Food and Agriculture, National Research and Innovation Agency of the Republic of Indonesia with grant number 6/III.11/HK/2024. The authors would like to thank Mr. Nikodemus Gultom, Mr. Taufik Yudhi and Mr. Nandang Jaenudin for their assistance in the roasting experiment.
Conflicts of interests. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
Declaration of generative AI and AI-assisted technologies in the writing process. During the preparation of this work the author(s) used ChatGPT (OpenAI Inc.) in order to improve grammar and readability of the manuscript. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

References

1. A. Rahayuningtyas, D. Sagita, M. A. Karim, L. E. Yulianti, D. P. Putri, I. F. Azizah, Y. H. Siregar, S. I. Kuala, D. D. Hidayat, and A. Darmawan, Unraveling the performance of fluidized coffee roasters under slow and fast roasting processes: energy, quality, and economic analysis, J. Biosyst. Eng. 49, 240 (2024). https://doi.org/10.1007/s42853-024-00230-3
2. J. Baggenstoss, L. Poisson, R. Kaegi, R. Perren, and F. Escher, Coffee Roasting and Aroma Formation: Application of Different Time−Temperature Conditions, J. Agric. Food Chem. 56, 5836 (2008). https://doi.org/10.1021/jf800327j
3. D. D. Hidayat, A. Rahayuningtyas, D. Sagita, M. Furqon, A. Sudaryanto, D. Gandara, T. Yudhi, and N. Jaenudin, Development and performance evaluation of a liquefied

petroleum gas-powered fluidized coffee roaster, Inmateh - Agricultural Engineering 69, 359 (2023)
4. D. Sagita, S. S. Mardjan, Suparlan, P. B. Purwandoko, and S. Widodo, Low-cost IoT-based multichannel spectral acquisition systems for roasted coffee beans evaluation: Case study of roasting degree classification using machine learning, Journal of Food Composition and Analysis 133, Article 106478 (2024). https://doi.org/10.1016/j.jfca.2024.106478
5. E. Bertone, A. Venturello, A. Giraudo, 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, 683 (2016). https://doi.org/https://doi.org/10.1016/j.foodcont.2015.06.055
6. F. de C. Pires, R. G. F. A. Pereira, M. R. Baqueta, P. Valderrama, and R. Alves da Rocha, Near-infrared spectroscopy and multivariate calibration as an alternative to the Agtron to predict roasting degrees in coffee beans and ground coffees, Food Chem. 365, Article 130471 (2021). https://doi.org/10.1016/j.foodchem.2021.130471
7. A. Praiphui, K. V. Lopin, and F. Kielar, Construction and evaluation of a low cost NIR-spectrometer for the determination of mango quality parameters, J. Food Meas. Charact. 17, 4125 (2023). https://doi.org/10.1007/s11694-023-01948-y
8. L. E. Yulianti, D. P. Putri, I. F. Azizah, S. Witman, M. A. Karim, and A. Rahayuningtyas, in BIO Web of Conferences, edited by A. Loch, R. C. Joshi, A. Munawar, and M. T. Multazam (EDP Sciences, Yogyakarta, 2023), p. 03014
9. J. Conzen, Multivariate Calibration: A Practical Guide for Developing Methods in the Quantitative Analytical Chemistry, 3rd ed. (Bruker Optik GmbH, Ettlingen, Germany, 2006)
10. D. Suhandy, M. Yulia, S. Widodo, H. Naito, and D. F. Al Riza, Fast authentication of Indonesian ground-roasted Arabica coffee adulterated with roasted soybean by portable LED-based fluorescence spectroscopy and chemometrics analysis, Food Chemistry 479, Article 143791 (2025). https://doi.org/10.1016/j.foodchem.2025.143791
11. A. M. de Souza Costa, K. Lirio Soares, L. de Souza Silveira, A. Carlos Verdin Filho, L. Louzada Pereira, V. Moreira Osório, M. Fronza, and R. Scherer, Influence of maturation and roasting on the quality and chemical composition of new conilon coffee cultivar by chemometrics, Food Research International 176, Article 113791 (2024). https://doi.org/10.1016/j.foodres.2023.113791
12. S. S. Mardjan, E. H. Purwanto, and G. Y. Pratama, Effect of Initial Temperature and Degree of Roasting on Physicochemical Properties and Taste of Solok Arabic Coffee, J. Keteknik. Pertan. 10, 108 (2022). https://doi.org/https://doi.org/10.19028/jtep.010.2.108-122
13. M. J. Lee, S. E. Kim, J. H. Kim, S. W. Lee, and D. M. Yeum, A study of coffee bean characteristics and coffee flavors in relation to roasting, Journal of the Korean Society of Food Science and Nutrition 42, 255 (2013). https://doi.org/10.3746/jkfn.2013.42.2.255
14. A. Tsiafitsa, V. Oikonomopoulou, M. Stramarkou, M. Krokida, and N. Papassiopi, Effect of heat treatment on physicochemical and sensory properties of selected coffee varieties, European Food Research and Technology 248, 2009 (2022). https://doi.org/10.1007/s00217-022-04025-8