Microchemical Journal 202 (2024) 110737

Contents lists available at ScienceDirect

Microchemical Journal

journal homepage: www.elsevier.com/locate/microc

Discrimination of Brazilian green canephora coffee beans by ultraviolet–visible spectroscopy as a non-target analysis: A tool for recognizing geographical indications

Venancio Ferreira de Moraes-Neto [a], Michel Rocha Baqueta [a], Elem Tamirys dos Santos Caramês [b], Felipe Bachion de Santana [c], Enrique Anastácio Alves [d], Juliana Azevedo Lima Pallone [a,*]

[a] University of Campinas – UNICAMP, School of Food Engineering, Department of Food Science and Nutrition, Campinas, São Paulo, Brazil
[b] Department of Microbiology, Institute of Biomedical Sciences, University of São Paulo, São Paulo, Brazil
[c] Agriculture and Food Development Authority-Teagasc, Wexford, Ireland
[d] Brazilian Agricultural Research Corporation – EMBRAPA Rondônia, Porto Velho, Rondônia, Brazil

ARTICLE INFO

Keywords:
Data science
Chemometrics
Discriminant analysis
Coffee beans

ABSTRACT

Specialty green coffee beans have a higher commercial value and some of them have recently been classified in Brazil based on the indication of provenance and denomination of origin. In this context, the classification of the type of coffee bean is still a challenge, using traditional analytical techniques. Thus, alternative analytical techniques, such as ultraviolet–visible spectroscopy (UV–Vis), as non-target analysis can be applied as a quick and reliable method of coffee classification using data science, such as chemometric tools. In the present study, UV–Vis were evaluated as a new strategy for discrimination of green beans of Brazilian specialty canephora coffees with recognized geographical indications (Robusta Amazônico and Conilon from state of Espírito Santo), for the first time. Spectra obtained from the aqueous extract of 222 samples. The Principal Component Analysis (PCA) was performed and subsequently Partial Least Squares with Discriminant Analysis (PLS-DA) model developed. The PCA indicated tendency to group the samples in their respective classes, pointing to the similarities in the spectra of samples of the same origin. The PLS-DA model obtained showed figures of merit values starting at 89.3% in the test set. The VIP scores showed that the variables associated with chlorogenic acids, caffeine and chlorophyll are the most important for differentiating the studied coffees. The results obtained showed that UV–Vis fingerprint – non-targeted analysis associated with PLS-DA is appropriate for the discrimination of green beans of Brazilian specialty coffee from different origins, in a simple way, using common equipment in several laboratories.

1. Introduction

The consumers increasingly demanding for better-quality coffees, specialty coffees have been conquering a wide and promising market, since they are recognized for their quality, historical aspects and production practices, based on environmental, social and economic sustainability [1]. The Arabica species (*Coffea arabica* L.) has always been more associated with fine beverages, while the canephora species (*Coffea canephora Pierre*), which used to be associated with lower quality, has recently been managed by some producers aiming to obtain specialty coffees [2,3].

In Brazil, the states of Espírito Santo, with a production of 11,221,000 bags, and Rondônia, with 2,263,100 bags in 2021, represent 82.76 % of all processed coffee in the country [4]. In Espírito Santo, the predominant cultivation is of the conilon botanical variety and in the state of Rondônia there is greater production of robustas, grown by indigenous peoples of the region. In 2021, the National Institute of Industrial Property (INPI) granted two geographical indications for these coffees [5]. The first geographical indication was the Indication of Provenance “Conilon Espírito Santo” and the second geographical indication was granted with the Denomination of Origin Matas de Rondônia, for Robustas Amazônicos coffees, produced by indigenous

* Corresponding author.
E-mail address: jpallone@unicamp.br (J.A.L. Pallone).

https://doi.org/10.1016/j.microc.2024.110737
Received 29 February 2024; Received in revised form 7 May 2024; Accepted 9 May 2024
Available online 10 May 2024

people and being the first denomination of origin for sustainable canephora coffees in the world. [6]. Specialty coffees have a higher commercial value and some of them have recently been classified in Brazil on the basis of indication of provenance and denomination of origin. In this context, mislabelling origin is characterised as fraud, but identifying the origin of coffee is still a challenge with the use of analytical techniques [7].

Recently, several studies had tested methods for the accurate discrimination of the origin of green coffee beans based on the chemical composition of samples. For instance, mineral profiling through Inductively Coupled Plasma Optical Emission Spectrometry (ICP-OES) and physicochemical composition analysis have been employed for the recognition of Arabica coffees in Brazil [8]; metabolomics based on Nuclear Magnetic Resonance (NMR) has been utilized for Arabica and Robusta coffees from Central America, South America, Africa, and Asia [9]; phenolic profiling via Ultra-High-Performance Liquid Chromatography coupled with Mass Spectrometry (UHPLC-MS) has been employed for Ethiopian Arabica coffees [10], and untargeted analysis using Liquid Chromatography coupled with Time-of-flight Mass Spectrometry (LC-TOF-MS) has been applied to Colombian Arabica coffees [11]. The results indicate that these methods are considered promising due to their achievements in classifications and predictions. However, most of the analytical strategies employed in these studies have relied on high-cost, labor-intensive, and time-consuming techniques and toxic residues generation [12].

Thus, there is great interest in developing efficient analytical techniques, based on the requirements of Industry 4.0 and in agreement to green analytical chemistry principles, which advocates simple, low-cost analysis capable of acquiring quick answers, during its production chain for this purpose and strong reduction or elimination of environmental negative impacts. Therefore, the development of faster and more environmentally friendly techniques to be used as tools to recognise specialty green canephoras coffees beans is still unheard of, thus guaranteeing consumers an authentic product and fair trade throughout the production chain [11,13].

In this sense, NIR spectroscopy has been successfully applied to the discrimination of coffees, due to all its characteristics related to green chemistry and the possibility of automation [14–16] But UV–Vis spectroscopy has been considered an effective technique and since it is widely used in routine analysis, the availability of UV–Vis spectrometers can be greater than most spectrometers [17]. In this scenario, ultraviolet–visible spectroscopy (UV–Vis) combined with data science by chemometric algorithms rises as an alternative method, which allow the development of simple, robust and economical option.

Chemometrics is the set of statistical, mathematical and graphical techniques for obtaining information based on chemical analysis data. Its use can be directed, for example: exploratory data analysis, by Principal Component Analysis (PCA) and sample classification, by Partial Least Squares Discriminant Analysis (PLS-DA), among others [18]. These two methods that are respectively supervised or unsupervised are applied as appropriate and promising tools for sample differentiation. [19,20]. PCA is based on a mathematical formulation that reduces the data dimension, enabling the recognition of patterns in the data set by highlighting their similarities and differences. PLS-DA, on the other hand, is defined as a tool that mathematically relates each analytical signal obtained from samples to a known class of samples of interest, by projecting the information contained in the spectra onto latent variables, where the maximization of covariance between the spectral data matrix and the class of target samples occurs [18].

The UV–Vis spectrophotometric fingerprinting (non-targeted) technique is used mainly for liquid samples or extracts, such as the studies reported for discrimination of roasted ground coffee as to type (caffeinated/decaffeinated) and conservation status (expired and not expired shelf life) using Successive Projection Modelling with Linear Discriminant Analysis (SPA-LDA) modelling [21]; identification of adulteration of ground roasted coffees by PLS-DA [22]; discrimination of Peaberry and traditional roasted coffees by PLS-DA [23] and PLS-DA to discriminate not only between species, but also between the geographical origins of green coffee [24]. These investigations show the versatility and effectiveness of discriminant analysis techniques in extracting information from UV–Vis data in order to categorize and discern various attributes in coffee samples.

For the discrimination of beans from specialty green canephora coffees with geographical indication, no published data were found in the scientific literature. Therefore, this study aims to evaluate the feasibility of using the UV–Vis spectroscopy technique associated with a PLS-DA model to discriminate Robusta Amazônico from Conilon coffees in the state of Espírito Santo, as a simple and low-cost method.

2. Material and methods

2.1. Samples and extraction process

Two types of green beans from Brazilian canephora speciality coffees with a geographical indication or recognized origin were assessed in this research: Robusta Amazônico (114 samples) and Conilon from the state of Espírito Santo (108 samples), totaling 222 samples. All the samples were provided by the Brazilian Agricultural Research Corporation (EMBRAPA), which collected them directly from the producers assuring the authenticity of the acquired samples. It is worth emphasizing that each sample came from a different producer.

The samples were grounded by freezing with liquid nitrogen and subsequently grinding in a home coffee grinder (Cadence, model MR35). The granulometry of the ground samples was standardized for all experiments using a 20 mesh/tyler − 0.85 mm sieve. Then, an aqueous extraction process was performed according to Souto [21] with adaptations: initially 1.0 g of each sample was placed on a paper filter in a glass funnel, then 10 mL of deionized water at 90–95 °C were poured over the samples and after calibrating to room temperature all extracts were diluted in distilled water in the ratio 1:500 (v/v). This dilution was defined by means of tests to find a ratio in which the maximum absorbances of the UV–Vis spectra reached values close to 1.

2.2. UV–VIS data acquisition

UV–VIS spectra of the diluted extracts were obtained using a multimode microplate reader (BMG Labtech, model Fluostar Omega, Germany), with a resolution of 1.0 nm. A quartz microplate at room temperature (25 ± 2 °C) was employed to collect the absorption spectra of each sample in the range 200 nm −1000 nm. The blank was evaluated using deionized water and subtracted to the diluted extracts spectra. A total of 222 UV–Vis spectra (1 spectrum per sample) with 781 variables per spectrum, were used for data analysis.

2.3. Exploratory analysis and model for classification of specialty canephora coffee beans with geographical indication by PLS-DA

MATLAB R2019a software (MathWorks IncNatick, USA) with PLS-toolbox 5.2 was used to apply data preprocessing techniques and apply chemometric tools to the UV–Vis spectra data.

All spectra were mean centred and the Principal Component Analysis (PCA) was applied, as an exploratory analysis, on the pre-processed spectra, to evaluate the distribution of samples according to spectral information, which could contribute to the grouping of samples. The number of Principal Components (PC) was set according to a cumulative explained variance and the smallest value of the Root Mean Square Error of Cross Validation (RMSECV) and the combination of the two most informative Principal Components (PCs) to aggrouping the samples according to the aim of this study presented.

For the classification of the coffees according to geographic origin and inherent botanical origin (Class 1 − Robusta Amazônico; and Class 2 − Conilon from the state of Espírito Santo), the PLS-DA model was developed with a training set, corresponding to 75 % of the samples and the remaining 25 % of the samples constituted the test set, the samples subdivision was performed by the Kennard-Stone algorithm [25]. Therefore, the training and test sets were composed of 167 samples and 55 samples, respectively. In addition, possible outliers were also checked in the model using leverage test and Q residuals. The optimal number of latent variables (LVs) was defined according to evaluation of the class boundary by Bayesian decision, the explained variance in Y and the ROC curves. Cross validation of the model was performed using a venetian blinds procedure with five samples.

The classification ability of the model was evaluated for each class by the sensitivity, specificity and non-error rate (NER) determined according to equations (1), 2 and 3, respectively. The correct discrimination rate (CDR) was also evaluated, taking into account all classes simultaneously (equation (4).Sensitivity refers to the ability to correctly predict all validation samples that do not belong to a given class. Specificity refers to the incorrect prediction of validation samples from other classes in a given class. NER refers to the rate of correctly classified samples in each class and CDR refers to the hit rate considering all classes. These values can take results from 0 % to 100 %, and have been calculated against the external validation data set. Models with figures of merit above 80 % in test are considered good predictors [7,18,26].

$$Sensibility(\%) = \frac{TP}{TP + FN}x100 \tag{1}$$

$$Specificity(\%) = \frac{TN}{TN + FP}x100 \tag{2}$$

$$NER(\%) = \frac{TP}{TP + FP + TN + FN}x100 \tag{3}$$

$$CDR(\%) = \frac{TP}{TP + FP + TN + FN}x100 \tag{4}$$

Where TP refers to true positives, FN to false negatives, TN to true negatives, and FP to false positives.

The regions of the spectrum (measured in wavelengths) that had the highest impact for modelling the classes were verified by calculating the projection importance variable (VIP). The variables that presented a VIP score value > 1 were considered more important (according to the wavelength ranges) [27].

3. Results and discussion

3.1. UV–VIS spectra obtained from canephora coffees bens with geographical indication and pre-processing methods

The spectra obtained from the 222 coffee samples did not show, in the region between 421 and 1000 nm, relevant spectral information as presented in Fig. 1A. The most informative band of the spectra is found in the UV region (220 nm – 420 nm), where useful transitions are $\pi \rightarrow \pi^*$ for compounds with conjugated double bonds, some $n \rightarrow \sigma^*$ transitions, and some $n \rightarrow \pi^*$ [28]. Absorption bands are seen in the 220 to 420 nm range. These absorption bands are mainly associated with the phenolic components present in green coffees [29].

In this case, the wavelength range that did not show relevant absorption bands were excluded in order to avoid negative interferences in the performance of the discriminant analysis models. Fig. 1B shows the average spectral profiles of the coffee grades in the working region between 220 and 420 nm that has the most information and the resulting spectra are suitable for visually differentiating between different coffee varieties.

3.2. Exploratory spectral analysis of coffee beans of different geographical indications: PCA

The PCA technique is generally used as an exploratory analysis before using supervised classification methods. Plotting the samples and their respective variables (spectral data obtained from UV–Vis) facilitates visualisation and grouping behaviour. The model for this study performed on the data after removing the non-informative region was established with the first two principal components (PC1 and PC2), which explain 97.62 % of the variance in the data. Fig. 2.A shows the results of PC1 and PC2.

The tendency to cluster the samples occurred along PC2, which alone explained 91.41 % of the variability in the data. The Conilon coffee samples from the state of Espírito Santo were mostly arranged in the positive quadrant of the scores, while the Robusta Amazônico samples were situated with negative scores. By analyzing the loadings graph (Fig. 2.B) it is possible to understand the reason for this separation. The spectral regions between 220–280 nm and 350–420 nm shows the highest and most positive loadings in PC2 and according to the literature, this is a region that is related to khaweol, cafestol, caffeine and trigonelline, so these are the compounds most associated with Conilon samples from the state of Espírito Santo. The region between 290 and 350 nm was relevant in the negative loadings of PC2 and is linked to Robusta Amazônico coffees and is a band linked to chlorogenic acids, especially caffeic, p-coumaric and ferulic acids [30].

Although PCA shows a clustering tendency, this type of modelling is not suitable for making predictions, and it is necessary to develop a supervised classification model to evaluate the possibility of using the UV–Vis spectra obtained for the discrimination of green coffee beans of different origins [26].

Previous to the supervised modelling, the presence of outliers was checked. There were no samples with high $Q_{residuals}$ and low leverage values. Since no samples had both high leverage and high $Q_{residuals}$, no outliers were detected in the training and test groups in any

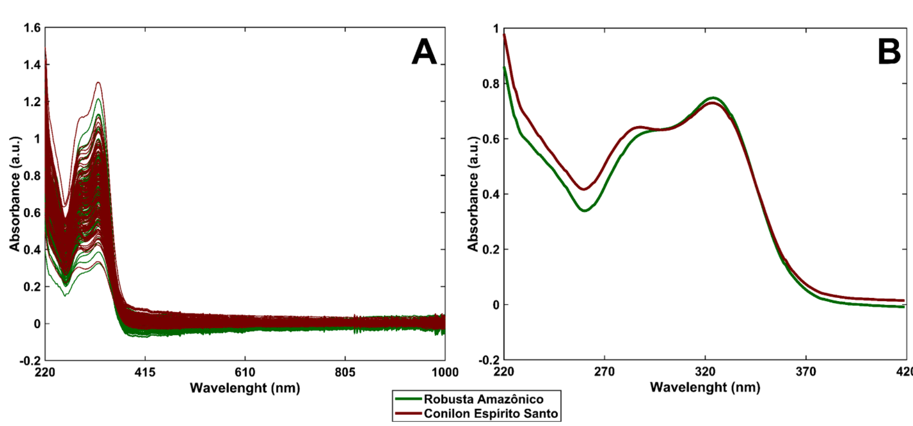

Fig. 1. Raw spectra (A) and average of the spectral region of interest (B).

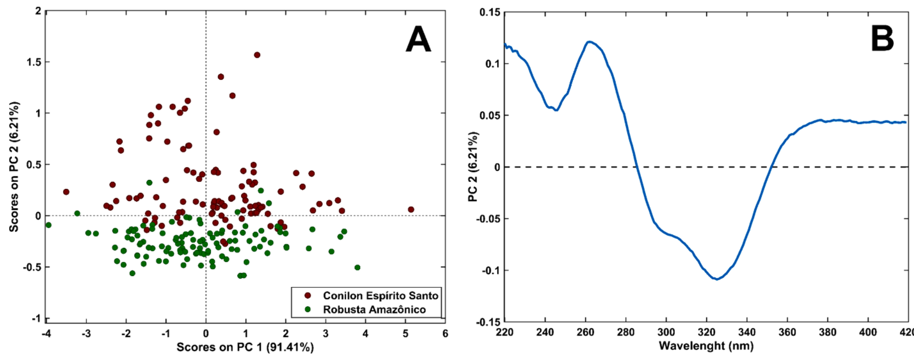

Fig. 2. PCA scores (A) and loadings (B).

preprocessing [31].

3.3. Models for discriminating samples of green coffee beans of different origins by PLS-DA

The confusion matrices and figures of merit of the PLS-DA model for discriminating between Robusta Amazônico and Conilon from the state of Espírito Santo are shown in Table 1. The number of latent variables (LVs) was set at 6, considering the class limit by Bayesian decision, thus reducing the chance of misclassification of the samples. The confusion matrices show the classification result in number of samples (true class versus predicted class) for each class in the training, cross-validation and test sets.

The analysis of the obtained model reveals significant figures of merit across different datasets and classes. Specifically, for the Robusta Amazônico class, the sensitivity values indicate strong predictive capability, with respective values of 94.2 %, 93.0 %, and 89.3.2 % for the training data, cross-validation, and test sets. Similarly, the specificity values for Robusta Amazônico demonstrate high accuracy, with values of 92.6 %, 92.6 %, and 100 % across the respective sets. Furthermore, the NER values exhibit consistent performance across all datasets, with scores of 94.2 %, 93.0 %, and 89.3 % for the training, cross-validation, and test sets, respectively.

For the Conilon from the state of Espírito Santo class, the obtained sensitivity values indicate scores of 92.6 %, 92.6 %, and 100 % across the respective datasets. Similarly, the specificity values for Conilon class samples demonstrate respective values of 94.2 %, 93.0 %, and 89.3 % across different datasets. Moreover, the NER values for Conilon samples highlight scores of 92.6 %, 92.6 %, and 100 % across the training, cross-validation, and test sets, respectively. The test set also revealed a CDR of 94.4 %, indicating the overall effectiveness of the model in accurately identifying instances across both classes. Therefore, the presented results demonstrate the robustness and effectiveness of the developed model in accurately classifying samples from both the Robusta Amazônico and Conilon from state of Espírito Santo. The figures of merit are within the range reported in the literature for coffee classification using UV–Vis spectroscopy combined with PLS-DA modelling, which ranged from 63.6 % to 100 % [23,24,32].

Since the UV–Vis spectra were obtained for aqueous extracts of green coffee beans, the technique was applied in its traditional form, using samples in solution after preparation of infusions, and showed feasibility. Fig. 3 shows the distribution performed for test samples. It is possible to identify that the grades of Robusta Amazônico and Conilon from state of Espírito Santo coffees were adequately separated and classified for most samples, since few samples were incorrectly predicted for the grades.

The VIP Scores (Fig. 4) indicate the most important variables for discrimination, which refer to the bands with greater differences among the coffee origins. The wavelength of highest intensity in each region delimited as VIP scores are considered as the main markers for the discrimination of coffee origins and can have their variabilities checked in boxplots (Fig. 5) for proof of contributions as PLS-DA predictors through the difference in analytical signal intensities.

The differences between the coffees are related to the band centred on 228 nm, which is associated with the composition of cafestol in the coffees, and the absorbances of this band, as can be seen in the boxplots in Fig. 5, were more intense for the Conilon coffees from the state of Espírito Santo. The band centred on 261 nm is a spectral region referring to caffeine and trigonelline; in this band the analytical signal was more intense for Conilon coffees from the state of Espírito Santo. In the region of 329 nm, is a region associated with chlorogenic acids and more expressive in Robusta Amazônico. The 417 nm wavelength is related to chlorophyll, has a higher value for the Conilon coffee class., indicating that there must be important differences in composition between the classes of coffees studied [29,33].

Among the compounds highlighted, the literature has proven that some are potential chemical markers of the geographical origin of coffees, especially trigonelline and chlorogenic acids, which are dependent on the attitude of the region of production. Caffeine and cafestol are not commonly reported as being dependent on the geographical origin of coffees, but in this study, they were shown to be a component of absorption in the VIP score region greater than 1, making them important for discrimination (Fig. 5) [30,34].

Robusta Amazônico green coffee beans from indigenous people were distinguishable from Conilon beans from the state of Espírito Santo using UV–Vis spectroscopy and PLS-DA supervised classification proved its efficiency for discrimination. With this, it was possible to discriminate the botanical/geographical origin of raw beans of Brazilian

Table 1
Confusion matrices and figures of merit of PLS-DA model.

LVs	Set	Geographical Indication	Confusion matrices		Figures of merit			
			Predicted as Conilon Espírito Santo	Predicted as Robusta Amazônico	Sens. (%)	Spec. (%)	NER (%)	CDR (%)
6	Training	Conilon Espírito Santo	75	6	92.6	94.2	92.6	93.4
		Robusta Amazônico	5	81	94.2	92.6	94.2	
	Cross-validation	Conilon Espírito Santo	75	6	92.6	93.0	92.6	92.8
		Robusta Amazônico	6	80	93.0	92.6	93.0	
	Test	Conilon Espírito Santo	26	0	100.0	89.3	100.0	94.4
		Robusta Amazônico	3	25	89.3	100.0	89.3	

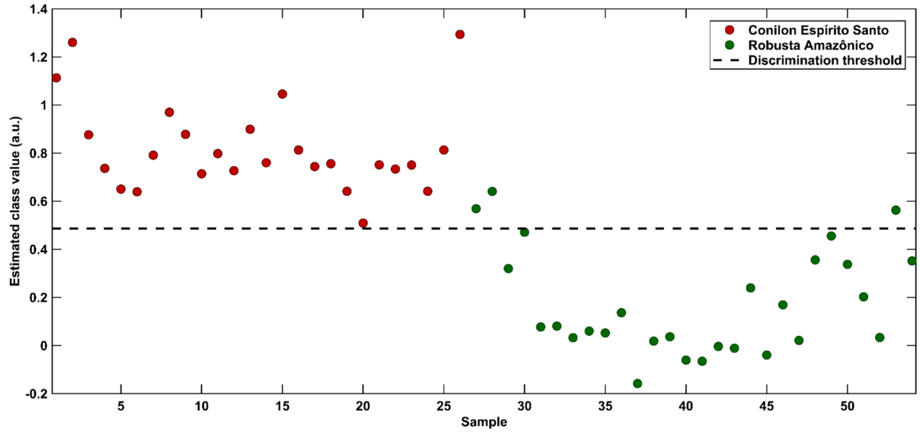

Fig. 3. Distribution of the test group in the PLS-DA model.

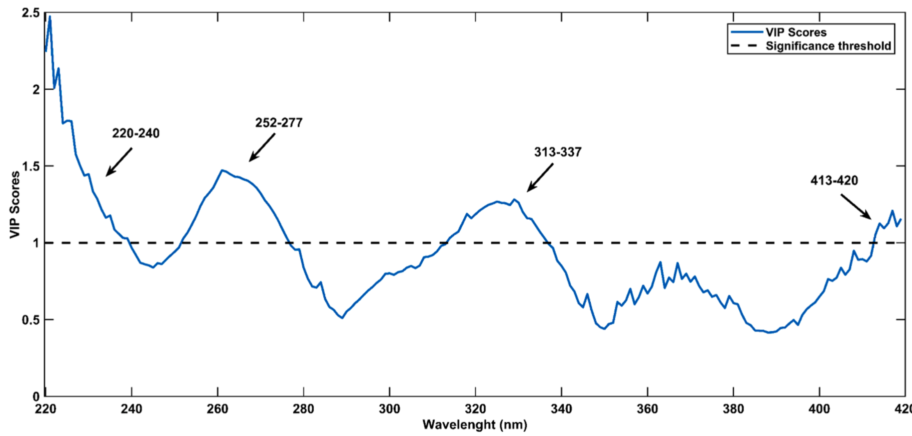

Fig. 4. VIP scores UV–Vis.

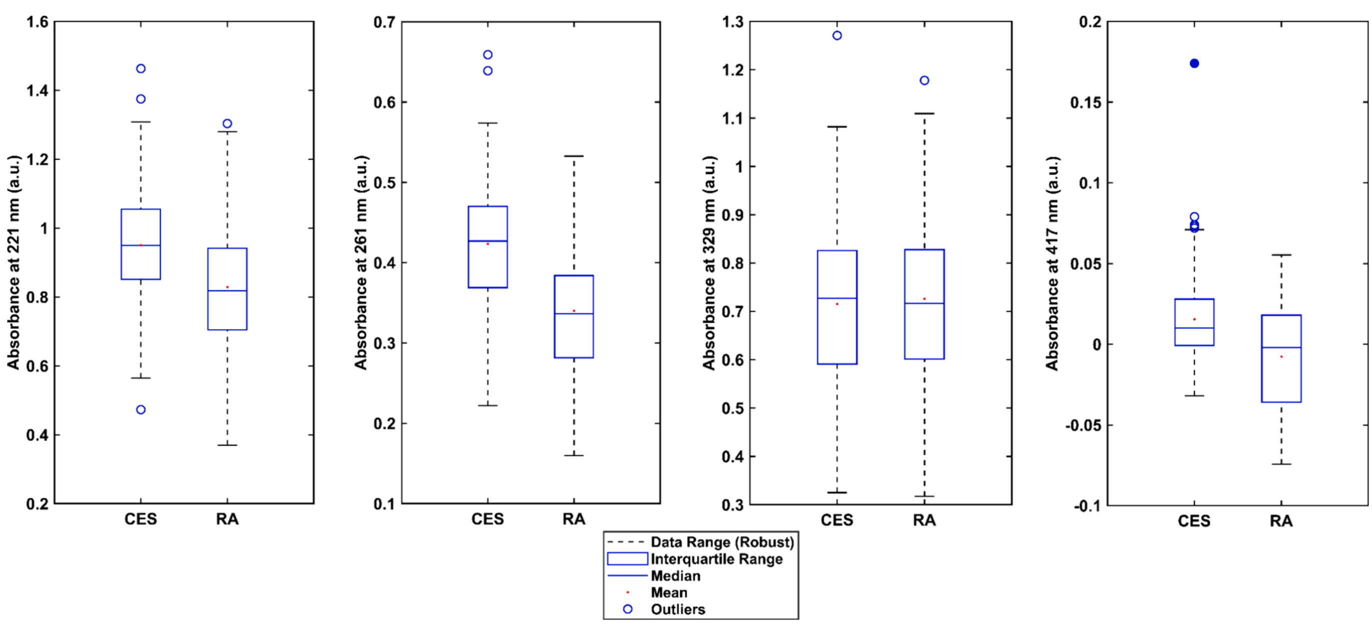

Fig. 5. Variability of the main VIP scores spectra of the classes Conilon from state of Espírito Santo (CES) and Robusta Amazônico (RA).

specialty canephora coffees with Geographical Indications, granted through UV–Vis spectroscopy from the diluted extract of the beans.

The model was considered satisfactory and has advantages for having simple sample preparation, no use of potentially toxic reagents and cheaper, thus being more feasible for the developed application, and can be used in quality control of producers, industries, certification agencies and associations involved in the evaluation of certified coffee beans [16] The application developed in this study also presents yet another alternative to the UV–Vis spectrometer, a piece of equipment commonly found in the food analytical laboratory sector, which is already widely used for other purposes to control the quality of coffees with univariate tests, such as the determination of caffeine, chlorogenic acids, antioxidant activity and total phenolics [35].

The results of this study are in line with those of previous studies that have used UV–Vis spectroscopy and discriminant analysis to analyse coffee. Souto et al. [22] used PLS-DA to detect adulteration in roasted and ground coffees, achieving remarkable figures of merit starting at 94 % and with SPA-LDA modelling achieving a 100 % correct classification rate. Suhandy and Yulia [23] employed PLS-DA to obtain 100 % correct discrimination in identifying Peaberry coffee from non-Peaberry coffee. Quan et al. [24], in their study of discriminating green coffees from various regions using PLS-DA, obtained sensitivity and specificity values starting at 63.6 %. These conclusions represent the effectiveness of UV–Vis spectroscopy associated with discriminant analysis, with specific variations in model performance observed in different applications and research contexts.

Marquetti et al. [14]and Giraudo et al. [16] utilized NIR spectroscopy for origin recognition, employing PLS-DA modelling and achieving correct classification rates of 35.7 % and 69.7 %, respectively. In a study by Mehari et al. [10], linear discriminant analysis (LDA) based on UHPLC-MS data resulted in an impressive recognition rate of 92 %. Furthermore, Endaye et al. [36] employed inductively coupled plasma-optical emission spectroscopy data, coupled with LDA, and achieved a commendable correct classification rate of 93.4 %.

4. Conclusion

The green beans of specialty coffee from different geographic/botanical origins (Robusta Amazônico and Conilon from the state of Espírito Santo) could be evaluated by UV–Vis spectroscopy and chemometrics, after obtaining aqueous infusions.

The exploratory PCA analysis indicated grouping tendencies of the samples in their respective classes (Robusta Amazônico and Conilon from the state of Espírito Santo), pointing to the similarities in the spectra of samples from different origins.

The UV–Vis spectral data were modelled with the PLS-DA method, enabling a method with a correct discrimination rate of 94.4 %. The VIP scores showed that the variables associated with chlorogenic acids, caffeine and chlorophyll are the most important for differentiating the studied coffees. The results obtained showed that UV–Vis spectroscopy associated with PLS-DA is appropriate for the classification of green beans of Brazilian specialty canephora coffee of different origins.

According to the results of this study, it is possible to affirm that the use of UV–Vis spectroscopy associated with chemometrics is an alternative and promising approach, characterized as a low cost, fast, safe, sensitive and green technique to classify green coffee beans from their aqueous extract, being an option for applications in quality control and origin certification procedures.

CRediT authorship contribution statement

Venancio Ferreira de Moraes-Neto: Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Michel Rocha Baqueta:** Methodology, Conceptualization, Writing – original draft. **Elem Tamirys dos Santos Caramês:** Writing – original draft, Methodology. **Felipe Bachion de Santana:** Writing – original draft, Methodology. **Enrique Anastácio Alves:** Resources. **Juliana Azevedo Lima Pallone:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

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

Data availability

The authors do not have permission to share data.

Acknowledgments

This work has been made possible through the work of many people and funding agencies. This work was supported by the National Council for Scientific Development (CNPq) (grant 133194/2020-5), São Paulo Research Foundation (FAPESP) (grant 2022/03268-3), and Coordination of Improvement of Higher Level Personnel (CAPES) – Finance Code 001. The authors would like to thank the EMBRAPA Rondônia and coffee producers for their invaluable help with the acquisition of the samples for the study.

References

[1] D.R. Reichman, Big Coffee in Brazil: Historical Origins and Implications for Anthropological Political Economy, Journal of Latin American and Caribbean Anthropology 23 (2018) 241–261, https://doi.org/10.1111/jlca.12293.

[2] P.S.M. Boaventura, C.C. Abdalla, C.L. Araújo, J.S. Arakelian, Value co-creation in the specialty coffee value chain: The third-wave coffee movement, RAE Revista De Administracao De Empresas 58 (2018) 254–266, https://doi.org/10.1590/S0034-759020180306.

[3] R.C.E. Dias, P. Valderrama, P.H. Março, M.B. dos Santos Scholz, M. Edelmann, C. Yeretzian, Quantitative assessment of specific defects in roasted ground coffee via infrared-photoacoustic spectroscopy, Food Chem. 255 (2018) 132–138, https://doi.org/10.1016/j.foodchem.2018.02.076.

[4] CONAB, Acompanhamento da Safra Brasileira de café - Safra, Primeiro levantamento, Obs. Agrícola. 10 (2023) 10–11. https://www.conab.gov.br/ info-agro/safras/cafe.

[5] Brazil, Brazilian coffees with geographical indication, (2021). https://www.gov.br/ agricultura/pt-br/assuntos/sustentabilidade/indicacao-geografica/arquivospublicacoes-ig/brazilian-coffees-with-geographical-indication.

[6] I.N.P.I. Brazil, Instituto Nacional da Propriedade Industrial, Indicações Geograficas (2021).

[7] C. Brodie, O. KraCht, A. hilKert, Tracing the geographical origin of roasted and green coffee using isotope fingerprints, J AOAC Int 102 (2019) 653–654, https://doi.org/10.5740/jaoacint.18-0314.

[8] D. Borsato, M.V.R. Pina, K.R. Spacino, M.B. dos S. Scholz, A.A. Filho, Application of artificial neural networks in the geographical identification of coffee samples, European Food Research and Technology 233 (2011) 533–543. https://doi.org/10.1007/s00217-011-1548-z.

[9] F. Wei, K. Furihata, M. Koda, F. Hu, R. Kato, T. Miyakawa, M. Tanokura, 13C NMR-based metabolomics for the classification of green coffee beans according to variety and origin, J. Agric. Food Chem. 60 (2012) 10118–10125, https://doi.org/10.1021/jf3033057.

[10] B. Mehari, M. Redi-Abshiro, B.S. Chandravanshi, S. Combrinck, M. Atlabachew, R. McCrindle, Profiling of phenolic compounds using UPLC-MS for determining the geographical origin of green coffee beans from Ethiopia, J. Food Compos. Anal. 45 (2016) 16–25, https://doi.org/10.1016/j.jfca.2015.09.006.

[11] D.E. Hoyos Ossa, R. Gil-Solsona, G.A. Peñuela, J.V. Sancho, F.J. Hernández, Assessment of protected designation of origin for Colombian coffees based on HRMS-based metabolomics, Food Chem 250 (2018) 89–97, https://doi.org/10.1016/j.foodchem.2018.01.038.

[12] J.A.L. Pallone, E.T. dos, S. Caramês, P.D. Alamar, Green analytical chemistry applied in food analysis: alternative techniques, Curr. Opin. Food Sci. 22 (2018) 115–121, https://doi.org/10.1016/j.cofs.2018.01.009.

[13] M.R. Malta, L. de, O. Fassio, G.R. Liska, G.R. Carvalho, A.A. Pereira, C.E. Botelho, V.P. Ferraz, A.D. Silva, A.W. Pedrosa, L.N. Alvaro, R.G.F.A. Pereira, Discrimination of genotypes coffee by chemical composition of the beans: Potential markers in natural coffees, Food Research International 134 (2020), https://doi.org/10.1016/j.foodres.2020.109219.

[14] I. Marquetti, J.V. Link, A.L.G. Lemes, M.B. dos, S. Scholz, P. Valderrama, E. Bona, Partial least square with discriminant analysis and near infrared spectroscopy for evaluation of geographic and genotypic origin of arabica coffee, Comput Electron Agric 121 (2016) 313–319, https://doi.org/10.1016/j.compag.2015.12.018.

[15] E. Bona, I. Marquetti, J.V. Link, G.Y.F. Makimori, V. da Costa Arca, A.L. Guimarães Lemes, J.M.G. Ferreira, M.B. dos Santos Scholz, P. Valderrama, R.J. Poppi, Support vector machines in tandem with infrared spectroscopy for geographical classification of green arabica coffee, LWT Food Sci. Technol. 76 (2017) 330–336, https://doi.org/10.1016/j.lwt.2016.04.048.

[16] A. Giraudo, S. Grassi, F. Savorani, G. Gavoci, E. Casiraghi, F. Geobaldo, Determination of the geographical origin of green coffee beans using NIR spectroscopy and multivariate data analysis, Food Control 99 (2019) 137–145, https://doi.org/10.1016/j.foodcont.2018.12.033.
[17] L.M. Reid, C.P. O'Donnell, G. Downey, Recent technological advances for the determination of food authenticity, Trends Food Sci. Technol. 17 (2006) 344–353, https://doi.org/10.1016/j.tifs.2006.01.006.
[18] M.M.C. Ferreira, Quimiometria: conceitos, métodos e aplicações, Editora da Unicamp, Campinas, SP, 2015.
[19] M. Efenberger-Szmechtyk, A. Nowak, D. Kregiel, Implementation of chemometrics in quality evaluation of food and beverages, Crit. Rev. Food Sci. Nutr. 58 (2018) 1747–1766, https://doi.org/10.1080/10408398.2016.1276883.
[20] D. Caballero, R. Ríos-Reina, J.M. Amigo, Chemometrics and Food Traceability, in: Comprehensive Foodomics, Elsevier, 2021: pp. 387–406. https://doi.org/10.1016/b978-0-08-100596-5.22859-x.
[21] U.T.C.P. Souto, M.J.C. Pontes, E.C. Silva, R.K.H. Galvão, M.C.U. Araújo, F.A. C. Sanches, F.A.S. Cunha, M.S.R. Oliveira, UV-Vis spectrometric classification of coffees by SPA-LDA, Food Chem. 119 (2010) 368–371, https://doi.org/10.1016/j.foodchem.2009.05.078.
[22] U.T.C.P. Souto, M.F. Barbosa, H.V. Dantas, A.S. Pontes, W. da, S. Lyra, P.H.G. D. Diniz, M.C.U. Araújo, E.C. Silva, Identification of adulteration in ground roasted coffees using UV-Vis spectroscopy and SPA-LDA, Lwt 63 (2015) 1037–1041, https://doi.org/10.1016/j.lwt.2015.04.003.
[23] D. Suhandy, M. Yulia, Peaberry coffee discrimination using UV-visible spectroscopy combined with SIMCA and PLS-DA, Int. J. Food Prop. 20 (2017) S331–S339, https://doi.org/10.1080/10942912.2017.1296861.
[24] N.M. Quan, H.M. Phung, L. Uyen, L.Q. Dat, L.G. Ngoc, N.M. Hoang, T.K.M. Tu, N. H. Dung, C.T.D. Ai, D.N.T. Trinh, Species and geographical origin authenticity of green coffee beans using UV–VIS spectroscopy and PLS–DA prediction model, Food Chemistry Advances 2 (2023), https://doi.org/10.1016/j.focha.2023.100281.
[25] R.W. Kennard, L.A. Stone, Computer Aided Design of Experiments, Technometrics 11 (1969) 137–148, https://doi.org/10.1080/00401706.1969.10490666.
[26] D. Ballabio, V. Consonni, Classification tools in chemistry. Part 1: Linear models, PLS-DA, Analytical Methods 5 (2013) 3790–3798, https://doi.org/10.1039/c3ay40582f.
[27] M.R. Baqueta, A. Coqueiro, P.H. Março, P. Valderrama, Multivariate classification for the direct determination of cup profile in coffee blends via handheld near-infrared spectroscopy, Talanta 222 (2021) 121526, https://doi.org/10.1016/j.talanta.2020.121526.
[28] P.H.G.D. Diniz, M.F. Barbosa, K.D.T. De Melo Milanez, M.F. Pistonesi, M.c.u. De Araújo, Using UV-Vis spectroscopy for simultaneous geographical and varietal classification of tea infusions simulating a home-made tea cup, Food Chem. 192 (2016) 374–379, https://doi.org/10.1016/j.foodchem.2015.07.022.
[29] G. Navarra, M. Moschetti, V. Guarrasi, M.R. Mangione, V. Militello, M. Leone, Simultaneous determination of caffeine and chlorogenic acids in green coffee by UV/Vis spectroscopy, J. Chem. 2017 (2017), https://doi.org/10.1155/2017/6435086.
[30] L.B. dos Santos, J. Tarabal, M.M. Sena, M.R. Almeida, UV-Vis spectroscopy and one-class modeling for the authentication of the geographical origin of green coffee beans from Cerrado Mineiro, Brazil, J. Food Compos. Anal. 123 (2023), https://doi.org/10.1016/j.jfca.2023.105555.
[31] A. Bagheri Garmarudi, M. Khanmohammadi, H. Ghafoori Fard, M. de la Guardia, Origin based classification of crude oils by infrared spectrometry and chemometrics, Fuel 236 (2019) 1093–1099, https://doi.org/10.1016/j.fuel.2018.09.013.
[32] M. Yulia, D. Suhandy, Indonesian palm civet coffee discrimination using UV-visible spectroscopy and several chemometrics methods, in, J Phys Conf Ser, Institute of Physics Publishing (2017), https://doi.org/10.1088/1742-6596/835/1/012010.
[33] A. Palacios-Morillo, Á. Alcázar, F. De Pablos, J.M. Jurado, Differentiation of tea varieties using UV–Vis spectra and pattern recognition techniques, Spectrochim. Acta A Mol. Biomol. Spectrosc. 103 (2013) 79–83, https://doi.org/10.1016/J.SAA.2012.10.052.
[34] G. de A. Mendes, M.A.L. de Oliveira, M.P. Rodarte, V. de Carvalho dos Anjos, M.J. V. Bell, Origin geographical classification of green coffee beans (Coffea arabica L.) produced in different regions of the Minas Gerais state by FT-MIR and chemometric, Curr Res Food Sci 5 (2022) 298–305. https://doi.org/10.1016/j.crfs.2022.01.017.
[35] AOAC, Official Methods of Analysis of AOAC International, 21st ed., AOAC International, 2019.
[36] M. Endaye, M. Atlabachew, B. Mehari, M. Alemayehu, D.A. Mengistu, B. Kerisew, Combining Multi-Element Analysis with Statistical Modeling for Tracing the Origin of Green Coffee Beans from Amhara Region, Ethiopia, Biol. Trace Elem. Res. 195 (2020) 669–678, https://doi.org/10.1007/s12011-019-01866-5.