WILEY | Hindawi

Research Article

Geographical Discrimination of Ground Amazon Cocoa by Near-Infrared Spectroscopy: Influence of Sample Preparation

Fabielle Negrão Ferreira [ID],[1] **Gilson Celso Albuquerque Chagas-Junior,**[1] **Mozaniel Santana de Oliveira** [ID],[2] **Jhonatas Rodrigues Barbosa,**[3] **Marcos Enê Chaves Oliveira,**[4] **and Alessandra Santos Lopes** [ID][1]

[1]*Laboratory of Biotechnological Process (LABIOTEC), Graduate Program in Food Science and Technology (PPGCTA), Federal University of Para (UFPA), Belém 66095-100, Pará, Brazil*
[2]*Laboratory AdolphoDucke, Research Field of the EmílioGoeldi Museum Paraense, Perimetral Avenue, 1901, Terra Firme, Belém, Pará, Brazil*
[3]*Laboratory for the Extraction of Plant Products (LABEX), Graduate Program in Food Science and Technology (PPGCTA), Federal University of Para (UFPA), Belém 66095-100, Pará, Brazil*
[4]*Embrapa EasternAmazon, Belém 66095-100, Pará, Brazil*

Correspondence should be addressed to Fabielle Negrão Ferreira; fabielle.negrao@gmail.com and Alessandra Santos Lopes; alessalopes@ufpa.br

This work presents the application of the NIR technique associated with exploratory analysis of spectral data by main principal components for the discrimination of Amazon cocoa ground seeds. Cocoa samples from different geographic regions of the state of Pará, Brazil (Medicilândia, Tucumã, and Tomé-Açu), were evaluated. The samples collected from each region were divided into four groups distinguished by the treatment applied to the samples, which were fermented (1-with fat and 2-fat-free) and unfermented (3-with moisture and 4-dried). Each set of samples was analyzed separately to identify the influence of moisture, fermentation, and fat on the geographical differentiation of the three regions. From the results obtained, it can be observed that it was not possible to differentiate the samples of seeds not fermented by geographic origin. However, fermentation was crucial for efficient discrimination, providing more defined clusters for each geographic region. The presence of fat in the seeds was a determinant to obtain the best model of geographic discrimination.

1. Introduction

Cocoa (*Theobroma cacao* L.) is one of the most important commercial crops in the world, contributing to the maintenance of inputs to various industries, such as food, pharmaceuticals, and cosmetics, contributing to the capital turnover of millions of dollars per year in addition to significantly influencing the world economy and income generation in several developing countries [1].

It is the primary raw material for the production of chocolate, which contains functional and bioactive groups, such as polyphenols, responsible for anticancer bioactivity [2], vasodilator [3], in addition to antioxidant bioactivity [4, 5], and anti-inflammatory [6].

The quality of cocoa seeds and various food products is correlated with several aspects, including genotype, crop conditions, and climate. Knowledge of geographical origin is often recognized and appreciated by the food industry and consumers and is an important factor that can add value to cocoa beans [7].

Unfermented cocoa seeds have an estimated chemical composition of 1% organic acids, 1% caffeine, 1-2% theobromine, 2-3% sucrose, 2-3% cellulose, 4–6% pentosans, 4–6% starch, 5 to 6% polyphenols, 10 to 15% proteins, 30 to

32% lipids, and 32 to 39% moisture [1, 8]. These compounds' profiles differ even before fermentation, according to the region where they were cultivated.

The fermentation and drying stages are crucial in cocoa processing, especially for the formation of color, aroma, and flavor of chocolate. These steps lead to profound changes in the chemical structure of cocoa seeds, changing the profile of phenolic, sugars, peptides, and triglycerides, among others [9–12]. Fat is the main constituent of fermented cocoa seeds composition and its fatty acid profile is changed during fermentation and directly affects the texture, viscosity, melting behavior, aroma, and flavor of the food produced from it [13].

According to the report from the Executive Committee of the Cocoa Plantation Plan (Ceplac), the state of Pará reinforced its position as the largest cocoa producer in Brazil in 2017 [14], and this production is dispersed in different localities, which produce fermented cocoa seeds with different sensory characteristics.

Geographical identification can help in the quality control and traceability management of these seeds. This identification and geographical authenticity can be performed through specific laboratory analyses, which are most time-consuming and tedious and require chemical products that are sometimes harmful to the environment [15–17].

The infrared (IR) spectroscopy technique, particularly in the Near-Infrared Region (NIR), is a simple, fast, nonchemical waste method and requires minimal sample handling. This method has been used effectively to determine the origin of various food and nonfood products [18].

Numerous food matrices have already been discriminated by their geographical origins using the NIR technique, among them, products such as honey [19], alcoholic beverages [20], mushrooms [21], green tea [22], butter [23], and wines [24].

The efficiency of the NIR technique has already been tested on cocoa from different regions of the world, and several components of the cocoa have been analyzed and correlated with the NIR method, such as color, volatile, phenolic compounds, antioxidants, fermentation index, and fat [7, 17, 25–29].

In addition, the multivariate classification methods (LDA, KNN, BPANN, and SVM) have also been tested by Teye et al. [7] for the discrimination of cacao originated in different regions. In this study, Teye et al. [7] used 194 samples and evaluated the grain discrimination from seven different regions (Ashanti, Brong-Ahafo, Central, Eastern, Volta, northwest, and southwest), according to the different growing regions in Ghana. PCA was used to test the models, as it provides relevant information about the trend of the samples and the formation of clusters [17, 25–29].

The purpose of this article was to use the NIR technique to evaluate the influence of sample preparation on the geographical discrimination of cocoa fruits in the cities of Medicilândia, Tucumã, and Tomé-Açu, which are the main important producing regions in the state of Pará, Brazil.

2. Material and Methods

2.1. Selecting Collection Regions. The region's selection for the execution of this present work was based on the production volume and the geographical distinction between them. Cities from the three main producing regions of the state of Pará (Medicilândia, Tucumã, and Tomé-Açu) and geographically distant from each other were chosen. In the *Transamazon region*, southwest of Pará, the city of Medicilândia (geographic coordinates 03°26′45″ S and 52°53′20″ W) was chosen, while in the southeastern and northeastern Pará, the cities of Tucumã (geographic coordinates 06°44′52″ S and 51°09′39″ W and Tomé-Açu (geographic coordinates 02°28′41.3″ S and 48°16′50.7″ W) were chosen, respectively. The production quality was the second criterion, where the region of Medicilândia stands out in the international market for having superior quality.

2.2. Sample Collection and Preparation. A total of 189 samples were used, 117 samples of raw cocoa fruits (RC) and 72 dried and ground fermented cocoa samples (DF). Samples were collected in three rural properties in the cities of Tomé-açu (RC = 40, DF = 20), Tucumã (RC = 40, DF = 21), and Medicilândia (RC = 37, DF = 31), corresponding to three different geographic regions of the state of Pará, Brazil. 15 samples were used for external prediction of the classification model performance.

Raw cocoa samples were collected following the traditional harvesting practice of each region, choosing random fruits at the edges and the center of the plantation.

After the harvest, the fruits were broken and stripped; the cotyledons were separated and milled individually in a multifunctional mill (model A11, IKA, Staufen, Germany). The milled material was then sieved in a 600 μm mesh and stored at −22°C until the analysis.

Fermented cocoa samples were obtained after the fermentation and drying processes, according to the traditional methods of each producer in the three regions mentioned above. These samples were dried to constant weight in an oven (model 80 series, Lucadema, São José do Rio Preto, SP, Brazil) at 80°C for moisture standardization. After the drying process, the dried fermented cocoa cotyledons were milled individually and then sieved in a 600 μm mesh and stored at −22°C until the analysis.

After the NIR analysis, the 117 samples of raw cocoa (RC) were dried to constant weight in an oven under air circulation at 80°C. After drying these samples, the spectra were again obtained.

2.3. Defatting of Cocoa Samples. Dried fermented cocoa samples (56 samples) were exhaustively degreased with petroleum ether (Synth, Diadema, SP, Brazil) at a temperature of 55 ± 1°C (method 963.15, AOAC), for 24 hours, according to the Association of Official Analytical Chemists-AOAC [30]. After removing the fat, DF samples were kept for 24 hours at 80°C in an air circulation oven (model 80 series, Lucadema, Brazil) to remove the solvent. After

solvent removal, dried fermented degreased cocoa samples (DFD) were stored at −22°C until analysis.

Through the procedures described above, it was possible to evaluate cocoa samples in 4 categories: raw cocoa (117), dried unfermented cocoa (117), dried fermented cocoa (72), and dried fermented degreased cocoa (56) (Table 1).

2.4. Obtaining NIR Spectra. NIR spectra were obtained using an MPA FT-NIR spectrometer (Bruker Optics, Ettlingen, Germany). The spectral data were acquired using absorbance mode in the spectral range from 3500 to $12500\,cm^{-1}$, with $16\,cm^{-1}$ resolution and an average of 32 scans per spectrum. For the samples spectral reading, vials with a volume of 3 ml were used with a capacity for about 1 g of the milled and sieved cotyledon. The spectral reading was performed at ±25°C.

2.5. Discrimination Model Development by Geographical Origin. The spectroscopy software used to construct discrimination models was OPUS 6.5 (Bruker Optics, Ettlingen, Germany). The spectral data were previously processed in OPUS 6.5, according to the type of sample. For the samples of cocoa RC, DF, and DFD, the vector normalization pretreatment (SNV- standard normal variate) was performed, while for the DU samples, SNV treatment was performed along with the first derivative.

After the application of pretreatments, the data were processed to develop the discrimination model by the geographical origin of each set of samples. The exploratory analysis method, PCA, was applied to multivariate data to make a visual inspection of results more evident.

The spectral region that has best differentiated the geographic regions was chosen by the software operator, based on information related to the sample chemical composition, such as water content, fat, and protein.

3. Results and Discussion

3.1. Spectral Analysis. The complete NIR spectra obtained from RC, DU, DF, and DFD samples are shown in Figure 1. The prominent peaks can be observed around 9000-8000, 8500-6000, 6000-5000, and $5000-4500\,cm^{-1}$ within the NIR range. In the RC spectrum (Figure 1(a)), two intense absorption peaks are observed, in 5312 and $7202\,cm^{-1}$, which are related to the water content in the samples, according to [7], corresponding to the region of the first Overton of the O-H stretch and ¯O-H deformation. The water absorption bands should be removed to reduce interference with the chemical structures corresponding to the group's CH and CO and combinations of amine groups [31]. With that, we could see in Figures 1(b) and 1(c) that after removing the water, the absorption peaks of other components in the samples became more evident.

The regions corresponding to 9000-7500, 6000-6030, 6030-4000, 4950-4500, and $4500-3850\,cm^{-1}$ were used for all the calibration model development.

None of the spectra presented in Figure 1 was possible to verify the differences between fruits from different analyzed

Table 1: Spectra collected by region and category: raw cocoa (RC), dried unfermented cocoa (DU), dried fermented cocoa (DF), and dried fermented degreased (DFD).

	RC	DF	DU	DFD
Tomé-Açú	40	20	40	19
Tucumã	40	21	40	19
Medicilândia	37	31	37	18
Total	117	72	117	56

regions, even in the fermented samples. These differences were only noticed after mathematical treatments through the application of statistical analysis techniques for model development.

3.2. Discrimination Models for Raw Cocoa. For the RC cocoa samples, it was impossible that the spectral band (6200-$4000\,cm^{-1}$) could discriminate groups of the fruits from different regions, as observed in Figure 2(a).

The graph of the scores presented in Figure 2(b), referring to the spectral range between 4000 and $6200\,cm^{-1}$, shows the lack of differentiation between the three regions' scores. Bands referring to the phenol group, which can be observed in the regions corresponding to $3562-3322\,cm^{-1}$, aromatic ring related bands, referring to the region of 2925-$2854\,cm^{-1}$ (attributed to the CH stretch of the aromatic ring) and $1645-1544\,cm^{-1}$ (attributed to C in the aromatic ring) [7], were overlapped. This effect is directly related to the amount of water and fats in the RC cocoa samples.

Therefore, high absorption of radiation NIR by water present in samples is a factor that contributes negatively to analysis, since the water spectra overlap with the other spectra of interest, making it challenging to construct an analytical and statistical model of geographic discrimination of the in-nature Amazonian cocoa.

3.3. Discrimination Models for Dried Unfermented Cocoa. After the removal of water by drying, the DU samples have shown a trend toward the formation of groups related to discrimination by geographic region. The best spectral range for the DU discrimination was between 4300 and $9300\,cm^{-1}$ observed in Figure 2(c).

These spectra located between 4300 and 4597, 4902–5199, 5805–6102, and $6406–9300\,cm^{-1}$ in the whole spectrum for DU are associated with the carbonyl group spectral regions, corresponding to stretching combination (CH_2 and CH), first harmonic of CH present in the aromatic ring, combination of C-C and C-N and second harmonic of N-H [15]. These molecular vibrations are caused by functional groups corresponding to polyphenols, alkaloids, vicilin class globulins, proteins, amines, acids, polysaccharides, and other aromatic compounds [7]. These diverse functional groups are the digital identity of each sample from different Amazon regions.

The differentiation between the samples from Tomé-Açu and Tucumã was evident, as observed in Figure 2(d). However, the scores related to the set of Medicilândia samples get mixed with the scores of the samples from the
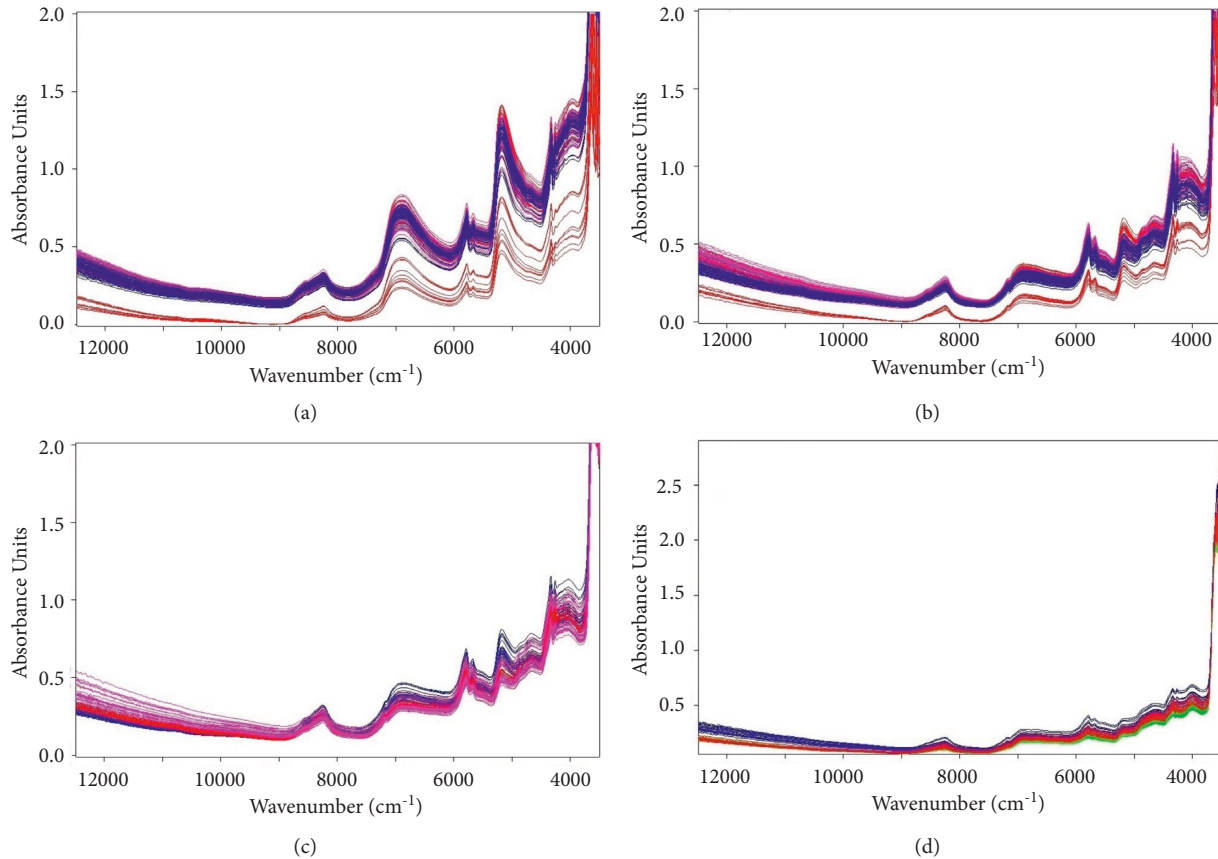
(a)

(b)

(c)

(d)

FIGURE 1: Absorption spectrum of (a) RC, (b) DU, (c) DF, and (d) DFD, in the range of 3600 to 12500 cm^{-1}.

other two regions. Several factors can explain the formation of characteristic groups of cocoa, such as chemical composition [32, 33], degree of fermentation [34], and genotype [35], which may be associated with the region of origin of the seeds [36].

According to [37], cocoa, like other food products, has its characteristics influenced by locality. This influence is directly related to the main components of alkaloid polyphenols, proteins, amines, polysaccharides, acids, and diverse aromatic compounds, as previously mentioned [7].

3.4. Discrimination Models for Fermented Cocoa. The fermentation of cocoa beans is essential, due to the diverse microbial processes developed as a consequence of changes in temperature, pH, and oxygen availability, promoting significant biochemical changes in the type and concentration of flavor precursors in cocoa beans [37] which is a crucial step in the formation of quality sensory attributes. Chemical and biochemical complex changes and interactions occur in the cocoa beans during drying, storage, and fermentation, contributing to the complexity and identity of the Amazon cocoa [38].

The fermentation produces flavor and aroma precursors, such as free amino acids and peptides from enzymatic degradation of proteins and sugars that reduce the enzymatic degradation of sucrose, as well as a significant increase

in volatile compounds such as organic acids, esters, alcohols, and aldehydes, after and during the fermentation of cocoa beans. The stoichiometry of their progenitors influences the concentration of these precursors; therefore, this is a response to the nutritional quality of the cultivated soil, from which the nutrients necessary for the synthesis of their parents were collected, revealing the good indicators of the production region [39].

Fermentation was a crucial factor in forming the characteristics of each region, because unlike the dried unfermented samples (DU), the dried fermented samples (DF), whose spectra are shown in Figure 2(e), allowed good discrimination between groups from different regions. The best spectral discrimination range observed was 6000 to 6800 cm^{-1}.

These results allow one to verify that there is a strong influence of the fermentation on the biochemical characteristics of the samples, to allow a clear distinction between them, as shown in Figure 2(f).

Caligiani et al. [10] reported that the progressive fermentation of cocoa seeds causes the hydrolysis of peptides in amino acids, reducing the violet color of the cotyledons at the end of the fermentation. The phenolic compounds are reduced upon drying; this reduction is mainly attributed to the enzymatic action, followed by nonenzymatic reactions due to quinone polymerization, with pH increase and high capitation of O_2 during sun drying [9].
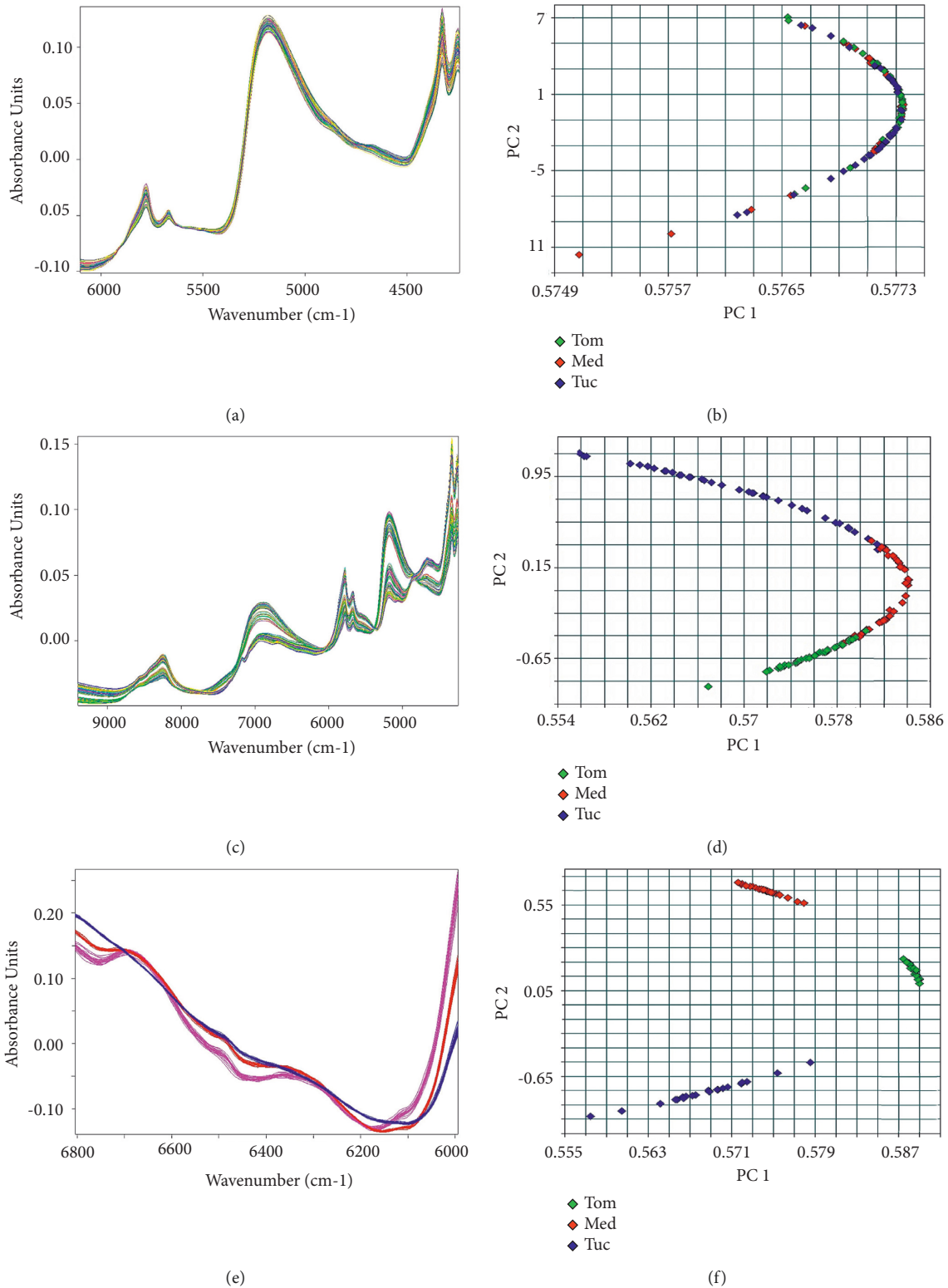
FIGURE 2: The absorption spectrum of the separation regions of (a) RC, (c) DU, and (e) DF, and their respective scores in the PCs graphs (b), (d), and (f).

The study of Sirbu et al. [12] estimated the increase of lipid content during fermentation from 3 to 6% when compared to the value before fermentation. In addition, the authors observed a change in the triacylglycerol profile caused by fermentation. Some polar triacylglycerols, such as derivatives of hydroxyl allyl fatty acid presented in
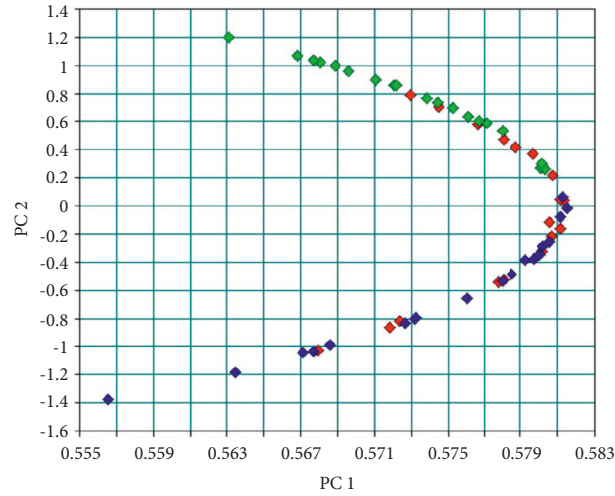
FIGURE 3: Scores of DFD cocoa samples.
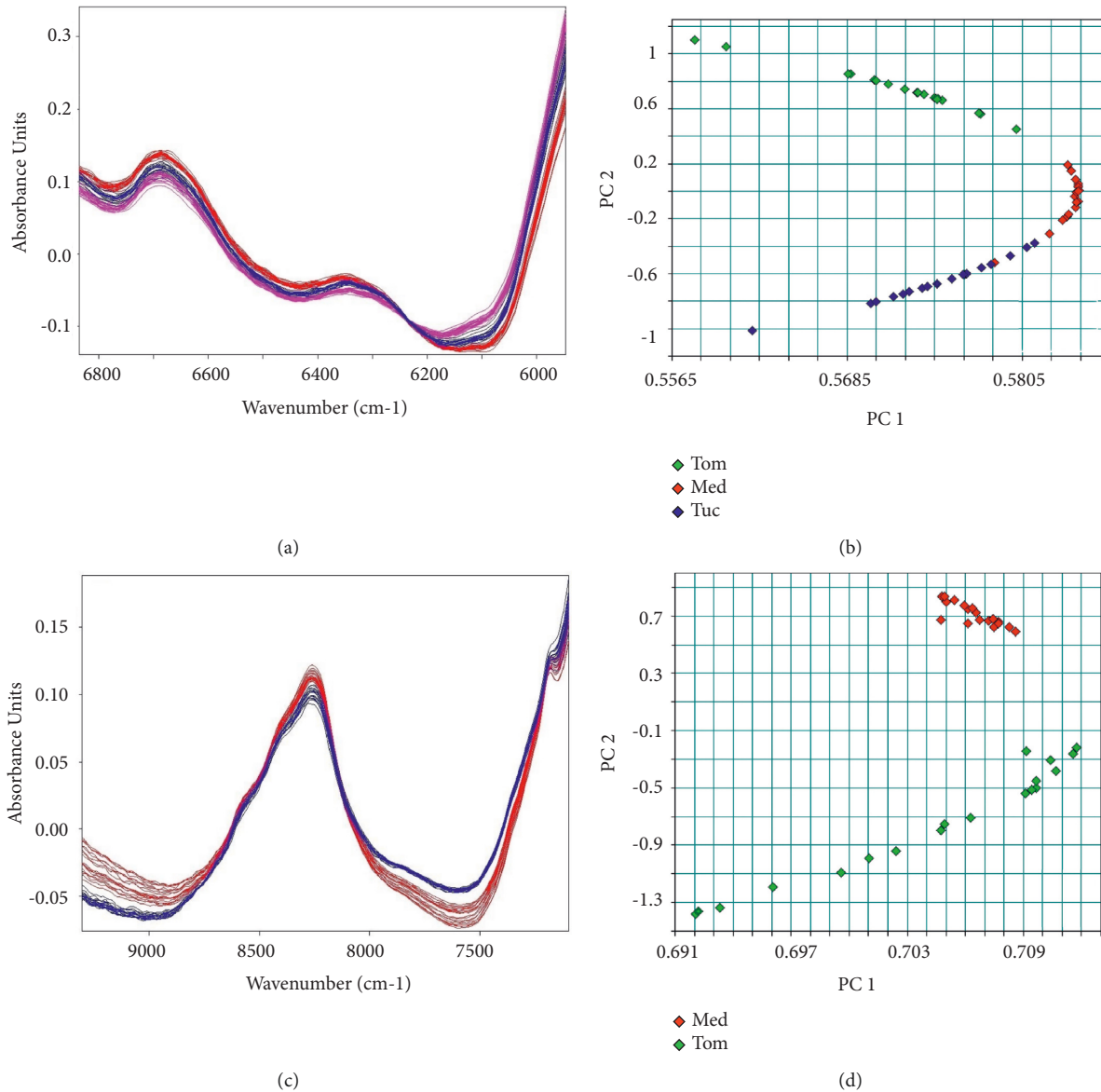

(a)

(b)

(c)

(d)

FIGURE 4: The absorption spectrum of the DFD separation regions (a, c) and their respective scores (b, d), for the two selected optimal separation regions.
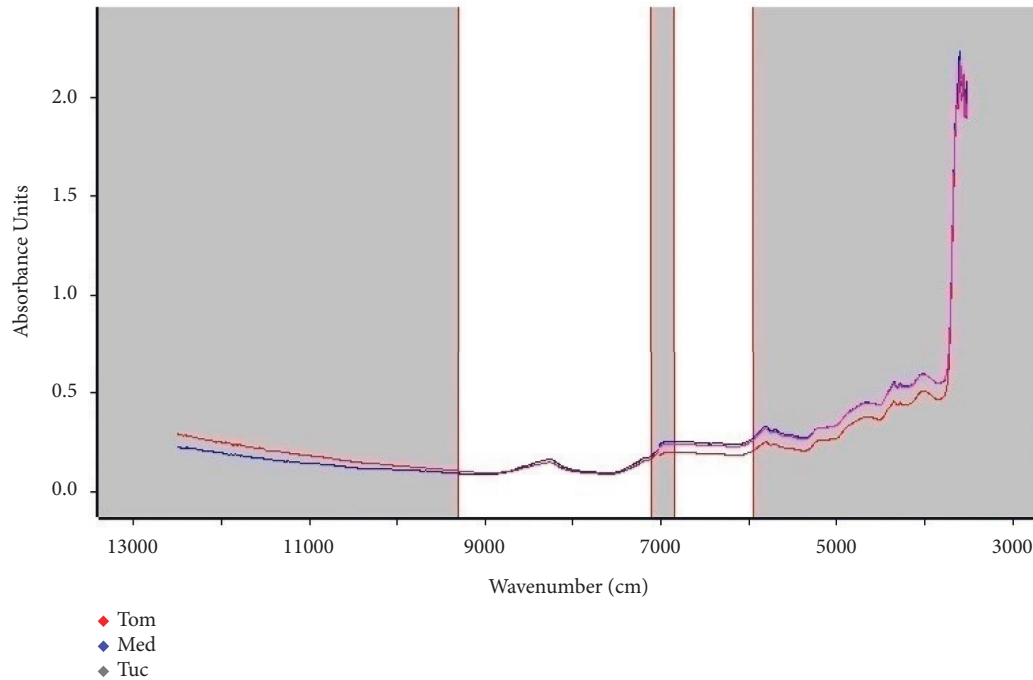
FIGURE 5: Absorption spectrum highlighting the two separation regions of DFD samples (1st and 2nd region).

unfermented cocoa seeds, were not found in dried fermented cocoa seeds, thus allowing chemical differentiation between unfermented and fermented cocoa seeds.

According to the literature, we can see that fermentation causes changes in the profile and concentration of various components of cocoa. Such components or concentrations of components can be characteristic for each region, and because of this, it is possible to carry out geographic identification. The NIR technique associated with chemometrics can previously identify the similarity and differences in various types of samples and classify them, making it possible to visually identify groups that have similar or distinct characteristics.

In this work, we used 15 external samples (not used in the construction of the model) to be able to test the efficiency of the constructed discrimination model (5 samples from Tomé-Açú, 5 from Tucumã, and 5 from Medicilândia). 100% of the samples were correctly classified according to their regions of origin.

3.5. Discrimination Models for Degreased Cocoa. To evaluate the importance of the lipid content of the samples in geographic differentiation, it was made the degreasing of the DF cocoa samples and the spectral evaluation was carried out again, using the same spectral treatment applied to discriminate the nondegreased fermented samples. Scores of the degreased samples are shown in Figure 3.

As observed in Figure 3, even though the mathematical model treatment was used for DF samples (dried fermented), it was not observed the separation between the three groups; however, a partial separation between two regions (Tomé-Açu and Tucumã) was observed, which was the same as visualized in the DU samples. This shows that the fat removal made it difficult to classify the samples in three different geographic

regions using the DFD sample discrimination model, validating partially the results of [40], who have shown a strong correlation of fatty acid composition in the geographic differentiation. In our research, we observed that it was possible to discriminate fermented samples more easily with their original butter constitution.

The lipids and proteins storage cells in cocoa seeds have complex cytology, composed of a compacted cytoplasm with multiple vacuoles, where proteins and lipids can be found, as well as other components such as starch granules, which are important for the definition of specific flavor and aroma characteristics. These biochemical and cytological characteristics of the cells change in percentage and morphological characteristics from region to region, affected by several physical, chemical, and climatic factors [39].

Since there was an indication of partial separation of DFD samples, it was developed a model from two separation steps. In the first separation step, the spectral region of 5950 to 6835 cm^{-1} was used (Figure 4(a)), in which the regions of Tomé-Açu and Tucumã were separated (Figure 4(b)). In the second step of the model, the sample data for the two groups that did not show clear separation in the first step (Medicilândia and Tucumã), the spectral region of 7100 and 9300 cm^{-1} was used (Figure 4(d)). Two spectral regions were used to build the complete separation model of the DFD samples, as shown in Figure 5.

The DFD cocoa seed sample separation model was performed in two stages. For the first stage, it was used the spectral data of the cocoa from the three regions (Medicilândia, Tucumã, and Tomé-Açu) (Figure 4(b)), and for the second stage, it was used the spectral data of the cocoa from Medicilândia and Tucumã, whose separation it was less evident in the first stage (Figure 4(c)). Thus, compared to the previous separation model (of DF seeds), we could observe

that the lipid fraction had considerable influence; however, it was not essential for geographic identification.

According to Sirbu et al. [12], after fermentation, cocoa presents a different profile of triacylglycerol and starch granules, which influence the spectroscopic behavior. And after the fat removal, the other components that changed during fermentation remain there, making geographic differentiation possible [12].

The application of the NIR technique is quite simple; however, the path to acquiring spectral information and analytical data is time-consuming and costly. Once the database of spectral and analytical information is obtained, the mathematical model can be constructed through the use of computational algorithms that plan and/or optimize such experimental procedures. From this, this elaborated model can be used to analyze external samples only with spectral information. The spectral data will be correlated with the data previously obtained and the result is estimated. In this way, the analysis becomes simple, fast, and economical. It is enough that, after the construction of the prediction model, the spectra of the material in question are obtained and the model is applied.

4. Conclusions

The results showed that it was possible to evaluate the influence of sample preparation on the geographic discrimination of cocoa from different geographic regions, in the Pará state, using the NIR technique.

The importance of the fermentation process for geographic identification was observed, and it was noticed provided specific characteristics for each region evaluated. The fermented cocoa lipids were crucial in the formation of groups in the fermented samples, also indicating that fermentation is a crucial step to develop specific characteristics for each collection region. Thus, it is assumed that the biotas of each region should affect the fatty acids contained in cocoa butter differently since similar results were not found in defatted samples.

The effect of the climate and production process, although not directly explicit, cannot be ruled out, considering that they can significantly influence the type of biota and the fermentation process, as well as other possible variables.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Ethical Approval

Ethics approval was not required for this research.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

References

[1] L. J. R. Lima, M. H. Almeida, M. J. Rob Nout, and M. H. Zwietering, "*Theobroma cacao* L., "The food of the gods": quality determinants of commercial cocoa beans, with particular reference to the impact of fermentation," *Critical Reviews in Food Science and Nutrition*, vol. 51, pp. 731–761, 2011.

[2] C. Oleaga, M. García, A. Solé, C. J. Ciudad, M. Izquierdo-Pulido, and V. Noé, "CYP1A1 is overexpressed upon incubation of breast cancer cells with a polyphenolic cocoa extract," *European Journal of Nutrition*, vol. 51, pp. 465–476, 2012.

[3] O. Khawaja, J. M. Gaziano, and L. Djoussé, "Chocolate and coronary heart disease: a systematic review," *Current Atherosclerosis Reports*, vol. 13, no. 6, pp. 447–452, 2011.

[4] V. T. T. Ho, J. Zhao, and G. Fleet, "The effect of lactic acid bacteria on cocoa bean fermentation," *International Journal of Food Microbiology*, vol. 205, pp. 54–67, 2015.

[5] A. Krähmer, A. Engel, D. Kadow et al., "Fast and neat—determination of biochemical quality parameters in cocoa using near infrared spectroscopy," *Food Chemistry*, vol. 181, pp. 152–159, 2015.

[6] S. A. Oyeleke, A. M. Ajayi, S. Umukoro, A. O. Aderibigbe, and O. G. Ademowo, "Anti-inflammatory activity of *Theobroma cacao* L. stem bark ethanol extract and its fractions in experimental models," *Journal of Ethnopharmacology*, vol. 222, pp. 239–248, 2018.

[7] E. Teye, X. Huang, H. Dai, and Q. Chen, "Rapid differentiation of Ghana cocoa beans by FT-NIR spectroscopy coupled with multivariate classification," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 114, pp. 183–189, 2013.

[8] J. E. Kongor, M. Hinneh, D. V. de Walle, E. O. Afoakwa, P. Boeckx, and K. Dewettinck, "Factors influencing quality variation in cocoa (*Theobroma cacao*) bean flavour profile—a review," *Food Research International*, vol. 82, pp. 44–52, 2016.

[9] J. F. M. Cruz, P. B. Leite, S. E. Soares, and E. S. Bispo, "Bioactive compounds in different cocoa (*Theobroma cacao*, L) cultivars during fermentation," *Food Science and Technology*, vol. 35, no. 2, pp. 279–284, 2015.

[10] A. Caligiani, A. Marseglia, B. Prandi, G. Palla, and S. Sforza, "Influence of fermentation level and geographical origin on cocoa bean oligopeptide pattern," *Food Chemistry*, vol. 211, pp. 431–439, 2016.

[11] R. Megías-Pérez, S. Grimbs, R. N. D'Souza, H. Bernaert, and N. Kuhnert, "Profiling, quantification and classification of cocoa beans based on chemometric analysis of carbohydrates using hydrophilic interaction liquid chromatography coupled to mass spectrometry," *Food Chemistry*, vol. 258, pp. 284–294, 2018.

[12] D. Sirbu, A. Grimbs, M. Corno, M. S. Ullrich, and N. Kuhnert, "Variation of triacylglycerol profiles in unfermented and dried fermented cocoa beans of different origins," *Food Research International*, vol. 111, pp. 361–370, 2018.

[13] E. Afoakwa, *Chocolate Science and Technology*, Wiley, Oxford, UK, 2010.

[14] IBGE, "Brazilian institute of geography and statistics, indicators IBGE. Agricultural production statistics," 2017, https://ww2.ibge.gov.br/home/estatistica/indicadores/agropecuaria/.

[15] A. Veselá, A. S. Barros, A. Synytsya, I. Delgadillo, J. Čopíková, and M. A. Coimbra, "Infrared spectroscopy and outer product analysis for quantification of fat, nitrogen, and moisture of cocoa powder," *Analytica Chimica Acta*, vol. 601, no. 1, pp. 77–86, 2007.

[16] P. A. da Costa Filho, "Rapid determination of sucrose in chocolate mass using near infrared spectroscopy," *Analytica Chimica Acta*, vol. 631, no. 2, pp. 206–211, 2009.

[17] E. Teye and X. Huang, "Novel prediction of total fat content in cocoa beans by FT-NIR spectroscopy based on effective spectral selection multivariate regression," *Food Analytical Methods*, vol. 8, no. 4, pp. 945–953, 2015.

[18] H. Büning-Pfaue, "NIR spectroscopy, a revolution in analysis," *New Food*, vol. 7, pp. 41–47, 2004.

[19] T. Woodcock, G. Downey, J. D. Kelly, and C. O'Donnell, "Geographical classification of honey samples by near-infrared spectroscopy: a feasibility study," *Journal of Agricultural and Food Chemistry*, vol. 55, no. 22, pp. 9128–9134, 2007.

[20] V. Uríčková and J. Sádecká, "Determination of geographical origin of alcoholic beverages using ultraviolet, visible and infrared spectroscopy: a review," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 148, pp. 131–137, 2015.

[21] Y. Chen, M. Y. Xie, Y. Yan et al., "Discrimination of Ganoderma lucidum according to geographical origin with near infrared diffuse reflectance spectroscopy and pattern recognition techniques," *Analytica Chimica Acta*, vol. 618, no. 2, pp. 121–130, 2008.

[22] Q. Chen, J. Zhao, and H. Lin, "Study on discrimination of Roast green tea (*Camellia sinensis* L.) according to geographical origin by FT-NIR spectroscopy and supervised pattern recognition," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 72, no. 4, pp. 845–850, 2009.

[23] M. Bassbasi, M. De Luca, G. Ioele, A. Oussama, and G. Ragno, "Prediction of the geographical origin of butters by partial least square discriminant analysis (PLS-DA) applied to infrared spectroscopy (FTIR) data," *Journal of Food Composition and Analysis*, vol. 33, no. 2, pp. 210–215, 2014.

[24] D. Cozzolino, W. U. Cynkar, N. Shah, and P. A. Smith, "Can spectroscopy geographically classify Sauvignon Blanc wines from Australia and New Zealand?" *Food Chemistry*, vol. 126, no. 2, pp. 673–678, 2011.

[25] A. M. C. Davies, J. G. Franklin, A. Grant, N. M. Griffiths, R. Shepherd, and G. R. Fenwick, "Prediction of chocolate quality from near-infrared spectroscopic measurements of the raw cocoa beans," *Vibrational Spectroscopy*, vol. 2, no. 2-3, pp. 161–172, 1991.

[26] P. C. Aculey, P. Snitkjaer, M. Owusu et al., "Ghanaian cocoa bean fermentation characterized by spectroscopic and chromatographic methods and chemometrics," *Journal of Food Science*, vol. 75, no. 6, pp. S300–S307, 2010.

[27] C. Hue, Z. Gunata, A. Bergounhou et al., "Near infrared spectroscopy as a new tool to determine cocoa fermentation levels through ammonia nitrogen quantification," *Food Chemistry*, vol. 148, pp. 240–245, 2014.

[28] E. Whitacre, J. O. Liver, R. Broek et al., "Predictive analysis of cocoa procyanidins using near-infrared spectroscopy techniques," *Journal of Food Science*, vol. 68, no. 9, pp. 2618–2622, 2003.

[29] N. Caporaso, M. B. Whitworth, M. S. Fowler, and I. D. Fisk, "Hyperspectral imaging for non-destructive prediction of fermentation index, polyphenol content and antioxidant activity in single cocoa beans," *Food Chemistry*, vol. 258, pp. 343–351, 2018.

[30] W. Horwitz and G. W. Latimer, *Official Methods of Analysis of AOAC International*, Association of Official Analytical Chemistry International, Rockville, MD, USA, 18th edition, 2006.

[31] S. Sunoj, C. Igathinathane, and R. Visvanathan, "Nondestructive determination of cocoa bean quality using FT-NIR spectroscopy," *Computers and Electronics in Agriculture*, vol. 124, pp. 234–242, 2016.

[32] D. F. Barbin, L. F. Maciel, C. H. V. Bazoni et al., "Classification and compositional characterization of different varieties of cocoa beans by near infrared spectroscopy and multivariate statistical analyses," *Journal of Food Science & Technology*, vol. 55, no. 7, pp. 2457–2466, 2018.

[33] E. Teye, X. Huang, L. K. Sam-Amoah et al., "Estimating cocoa bean parameters by FT-NIRS and chemometrics analysis," *Food Chemistry*, vol. 176, pp. 403–410, 2015.

[34] E. K. Anyidoho, E. Teye, and R. Agbemafle, "Nondestructive authentication of the regional and geographical origin of cocoa beans by using a handheld NIR spectrometer and multivariate algorithm," *Analytical Methods*, vol. 12, no. 33, pp. 4150–4158, 2020.

[35] J. P. Cruz-Tirado, J. A. Fernández Pierna, H. Rogez, D. F. Barbin, and V. Baeten, "Authentication of cocoa (Theobroma cacao) bean hybrids by NIR-hyperspectral imaging and chemometrics," *Food Control*, vol. 118, Article ID 107445, 2020.

[36] A. C. Oliveira, A. Marien, J. Hulin et al., "Development of real-time PCR methods for cocoa authentication in processed cocoa-derived products," *Food Control*, vol. 131, Article ID 108414, 2022.

[37] D. Kadow, J. Bohlmann, W. Phillips, and R. Lieberei, "Identification of main fine or flavour components in two genotypes of the cocoa tree (*Theobroma cacao* L.)," *Journal of Applied Botany and Food Quality*, vol. 86, pp. 90–98, 2013.

[38] V. T. T. Ho, J. Zhao, and G. Fleet, "Yeasts are essential for cocoa bean fermentation," *International Journal of Food Microbiology*, vol. 174, pp. 72–87, 2014.

[39] E. O. Afoakwa, J. E. Kongor, J. Takrama, and A. S. Budu, "Changes in nib acidification and biochemical composition during fermentation of pulp pre-conditioned cocoa (*Theobroma cacao*) beans," *International Food Research Journal*, vol. 20, pp. 1843–1853, 2013.

[40] M. Torres-Moreno, E. Torrescasana, J. Salas-Salvadó, and C. Blanch, "Nutritional composition and fatty acids profile in cocoa beans and chocolates with different geographical origin and processing conditions," *Food Chemistry*, vol. 166, pp. 125–132, 2015.