Agrotrópica 36(1,2): 7 - 18.
Centro de Pesquisas do Cacau, Ilhéus, Bahia, Brasil

QUALITY RANGE FOR BIOCHEMICAL ATTRIBUTES OF CACAO BEANS

Guilherme Amorim Homem de Abreu Loureiro[1], Quintino Reis de Araujo[2], Dario Ahnert[3], Etztli Itzel Morales Reyes[4], Rodrigo Antonio Escalona Valdez[5], Virupax Chanabasappa Baligar[6], Yuncong C. Li[7]

[1]Postgraduate Program in Forest Science, Postgraduate College (Postdoctoral Stays - Conahcyt), Campus Montecillo, Texcoco 56230, State of Mexico, Mexico. E-mail: gahal.85@gmail.com; [2]Department of Agricultural and Environmental Sciences at the State University of Santa Cruz (UESC), Jorge Amado Road km 16 45662-000, Ilhéus, Bahia, Brazil (Retired). E-mail: quintino@uesc.br; [3]Department of Biological Sciences at the UESC, Jorge Amado Road km 16 45662-000, Ilhéus, Bahia, Brazil. E-mail: darioa@uesc.br; [4]Departament of Hydrosciences at the Postgraduate College, Montecillo Campus, km 36.5 México 136 5, Montecillo, 56230 Texcoco, State of Mexico. E-mail: etizel114@hotmail.com; [5]Superior Normal School of Mexico (ENSM), Azcapotzalco, Manuel Salazar 201, Hacienda del Rosario, Ex el Rosario, Azcapotzalco, 02420, Ciudad de México, México. E-mail: rodrigo.escalona@aefcm.gob.mx; [6]US Department of Agriculture (USDA), Agricultural Research Service (ARS), Beltsville Agricultural Research Centre, Beltsville 20705-2350, Maryland, United States (Retired). E-mail: baligarvc22@gmail.com; [7]Department of Soil, Water, and Ecosystem Sciences, Tropical Research and Education Center. Institute of Food and Agriculture Sciences, University of Florida, Homestead 33031, Florida, United States. E-mail: yunli@ufl.edu.

This research aimed to establish critical limits for the commercial classification and biochemical attributes of cacao beans that would reflect an optimal quality range by classifying the observations of the samples by the fermentation index (FI). Boundary line regressions represented the response of brown cacao beans, purple cacao beans, partially brown cacao beans, acetic acid, lactic acid, sucrose, fructose, glucose, total protein, total free amino acids, total lipids, theobromine, caffeine, (+)-catechin, (-)-epicatechin, and total phenolic substances as a function of pH variations sensitive to prior fermentation. The elliptical clustering was used to delimit a normal probability of occurrence of sample observations classified with FI values between 1 and 1.4. The optimal fermentation condition for these attributes is in the pH range between 5.79 and 6.05. Purple cacao beans, lactic acid, and sucrose were the most sensitive attributes to the FI quality range.

Key words: *Theobroma cacao* L., post-harvest processing, cacao quality, multivariate analysis, biochemical regressions, elliptical clustering.

Faixa de qualidade de atributos bioquímicos de amêndoas secas de cacau. O objetivo desta pesquisa foi estabelecer limites críticos para os atributos da classificação comercial e bioquímicos das amêndoas de cacau que refletissem uma faixa ótima de qualidade classificando as observações das amostras pelo índice de fermentação (IF). As regressões da linha de fronteira representaram a resposta de amêndoas de cacau marrons, amêndoas violetas, amêndoas parcialmente marrons, ácido acético, ácido lático, sacarose, frutose, glicose, proteína total, aminoácidos livres totais, lipídios totais, teobromina, cafeína, (+) -catequina, (-)-epicatequina e substâncias fenólicas totais em função das variações de pH sensíveis à fermentação prévia. O agrupamento elíptico foi utilizado para delimitar uma faixa de probabilidade normal de ocorrência de observações amostrais classificadas com valores de IF entre 1 e 1,4. A condição ótima de fermentação para estes atributos é encontrada na faixa de pH entre 5,79 e 6,05. Amêndoas de cacau violetas, ácido lático e sacarose foram os atributos mais sensíveis à faixa de qualidade IF.

Palavras-chave: *Theobroma cacao* L., processamento pós-colheita, qualidade do cacau, análise multivariada, regressões bioquímicas, agrupamento elíptico.

Recebido para publicação em 19 de julho de 2024. Aceito em 27 de junho de 2024
DOI:10.21757/0103-3816.2024v36n1,2p7-18

Introduction

The cacao (*Theobroma cacao* L.) production and its food quality are largely influenced by environmental, technical, and social factors (Araujo, Q.R. de et al., 2017, 2021; Araujo et al., 2018; Araujo et al., 2014, 2018; Franzen and Borgerhoff Mulder, 2007; Loureiro et al., 2016, 2017; Ploetz, 2016) for obtaining the biochemical profile of cacao beans that meets the criteria for industrial use and of food safety (Araujo et al., 2019; Cinar et al., 2021; Loureiro et al., 2017). For this reason, food science has sought to characterize and analyse the processes that define the quality attributes of dry cacao beans (Cinar et al., 2021; Loureiro et al., 2017). Cacao Quality Indexes (CQIs) were proposed recently to integrate groups of biochemical attributes according to the different interests of this primary product, such as the development of flavour and beneficial properties for human health (Araujo et al., 2019; Araujo et al., 2014). However, even with the literature support (Araujo et al., 2019; Araujo et al., 2014; Loureiro et al., 2017), the definition of the critical limits of the attributes used in these indexes needs to be revised in the biochemical context of the post-harvest processes of fermentation and drying.

The quality of dry cacao beans (DCB), the raw material for making chocolate results from biochemical transformations that occur in the seeds beginning in the fermentation process and ending in the drying process. Therefore, the quality of beans must be defined from thresholds (quality ranges) established experimentally for specific attributes, through monitoring and control of fermentation and drying conditions. However, most of these thresholds in the literature lack an analytical context because they are derived from samples that omit the origin, genotype of the cacao, and post-harvest processing conditions (Araujo et al., 2019; Loureiro et al., 2017). Therefore, the establishment of limits without due scientific basis may overestimate or underestimate the quality of commercial lots of cacao beans (MAPA, 2008), whose production and processing conditions are equally unknown.

The fermentation quality of Forastero or Trinitarian cacao can be monitored by the anthocyanins breakdown and formation of by-products such as cyanidin-3-b-D-galactoside and cyanidin-3-a-L-arabinoside (Gourieva and Tserevitinov, 1979; Kim and Keeney, 1983). The oxidation of these products is related to the formation of brown pigments in completely fermented beans, which is detected by spectrophotometry from the increase in absorbance at 460 nm, and the decrease in absorbance at 530 nm (bin Said, 1989; Gourieva and Tserevitinov, 1979; Shahrir and Dimick, 1986). The absorbance ratio of 460 nm/530 nm is taken as the Fermentation Index (FI) (Gourieva and Tserevitinov, 1979). The literature indicated an FI quality range between 1 and 1.4 (Khairul Bariah, 2014; Romero-Cortes et al., 2013). The under-fermentation range corresponded to values lower than 1, and the over-fermentation range corresponds to values higher than approximately 1.4. Furthermore, an inverse relationship between the Fermentation Index (FI) and the potential of hydrogen (pH) was expected, which can be explained by the increase in total acidity due to the formation of acidic compounds during the fermentation process (Ilangantileke et al., 1991; Apriyanto et al., 2016). Therefore, from fermentation to drying, pH is a fundamental attribute to monitor the beans' acidity and is widely used by the cacao industry (Araujo et al., 2019; Loureiro et al., 2017).

Loureiro et al. (2023) implemented some techniques for the graphical analysis of free amino acid concentrations using FI and pH as predictive variables. For an adequate interpretation of results, certain particularities of the graphic analysis method need to be considered. The FI cannot be determined on DCB. Only pH or total acidity are variables capable of indicating the fermentation degree of sample observations. Although FI cannot be determined in DCB samples, the quality range ($1 < FI < 1.4$) can be detected and monitored by classifying cacao samples. These sorted samples have their distribution and probability domain. Therefore, it is also possible to verify clusters of observations of 'best fermented' samples within the two-dimensional space of the response of a biochemical attribute to pH variability. To verify the response trend of biometric or biochemical variables as a function of a predictor variable, that is, whose variability can be controlled or induced (by humans or by the environment), the polynomial borderline regression technique can be used (Walworth et al., 1986; Webb, 1972, 1973).

Retrospective fermentation monitoring by FI was carried out by clustering subsamples (Loureiro et al., 2023). Elliptical geometry is commonly used for clustering subsamples in the multivariate space generated by covariance or correlation matrices (Friendly et al., 2013; Jolicoeur, 1959). Elliptical clusters are calculated with the average as the centre of the subsample (Jolicoeur, 1959). The ellipse axis and its directions are calculated based on the maximum variation of the subsamples represented by the standard deviations of both axes (Friendly et al., 2013). Observations of samples sorted by FI values allowed an elliptical area with a multivariate normal distribution to be plotted on the scatterplot as the highest probability domain of the best fermentation condition. Thus, this research aimed to calibrate two-dimensional ranges of the response of biochemical attributes of DCB to pH variations through elliptical clustering of sample observations classified by ideal FI values.

Material and Methods

Cacao origin, sampling, and post-harvest processes

The fermented cacao beans (FCB) and DCB of the Porto Híbrido 16 (PH-16) clone, from 12 cropping sites in the south-eastern region of Bahia, are the origin of the analysed material. The trial was conducted in a completely randomized design with 36 observations (samples), from 12 cultivation sites and three replications by location. Fruit collections were carried out in November 2008, during the region's second annual cacao harvest. Fermentation processes of 168 h (7 days) and drying of 192 h (8 days) are completely reported by Loureiro, G.A.H. de A. et al. (2016, 2017).

Fermentation index

The FI determination is based on the method adapted from Gourieva and Tserevitinov (1979). Samples of 0.5 g of milled FCB were homogenized with 50 ml of a mixture of CH_4O (methanol) and HCl (hydrochloric acid) in a 97:3 volume ratio. The homogenate was left in a refrigerator at 8°C for 20 hours and then vacuum filtered. The FI, characterized by the percentage of beans that underwent different levels of fermentation indicated by their colour (completely brown, partially brown, violet, and slate) in each sample, was obtained by calculating the ratio between the absorbance at 460 nm and the absorbance at 530 nm on the spectrophotometer (FEMTO® 600 plus, FEMTO, Brazil).

pH

The hydrogen ionic potential (pH) of the DCB was determined by method 970.21-1974 of the Association of Official Analytical Chemists (AOAC) International (2005a).

Longitudinal cut for commercial classification

The longitudinal cut of DCB for commercial classification was determined according to Normative Instruction 38/2008 of the Ministry of Agriculture, Livestock and Food Supply (MAPA) from Brazil.

Organic acids and simple carbohydrates

The contents of organic acids (acetic acid - AA and lactic acid - LA) and simple carbohydrates (glucose - GLC, fructose - FRU, and sucrose - SUC) in DCB were determined by High-Performance Liquid Chromatography (HPLC) according to the method described by Schwan and Souza (1986), modified for this study. The samples triplicate of 20 g of milled DCB (endosperm) were weighed in a 300 ml beaker to which 200 ml of distilled water was added, then homogenised and centrifuged at 3000 revolutions per minute (rpm) for 10 minutes at 20 °C. Approximately 4 ml of the supernatant was placed in 2 Eppendorf Tubes® and centrifuged again at 14,000 rpm at 4°C for 10 minutes. The extracts were filtered through a 0.45 ìm polyvinylidene fluoride (PVDF) membrane filter, and an aliquot (20 µl) was injected into the HPLC system (model-302, Gilson, Middleton, WI) equipped with a refractive index detector (RID-300, Japan Spectroscopic Co., Ltd., Tokyo). The separation of organic acids and simple carbohydrates was performed on an Aminex Ion Exclusion HPX-87H column (300 x 7.8 mm) using 0.005 N sulfuric acid as the eluent. Peak acquisition and integration were conducted using Star Chromatography Workstation 6.0 software (©Varian, Inc. 2004, Walnut Creek, CA, USA), based on previously prepared standard curves, with HPLC grade standards (Sigma-Aldrich, Merck, St Louis, MO, USA).

Protein and total free amino acids

Total protein (PRO) content in DCB (endosperm) was estimated based on conversion factor of 6.25 of the total nitrogen, which was determined using the micro-Kjeldahl assay according to the method 31.1.08 (AOAC International, 2005a). The steps for the analysis and determination of free amino acid (FAA) contents in DCB by HPLC were described by Loureiro et al. (2023).

Total lipids

The total lipid (LIP) content of DCB was determined according to the modified method 963.15-1973 of AOAC International (2005b). Samples of 5 g of milled DCB (endosperm) were placed in a Soxhlet extractor (including a heating plate and circulator with refrigeration). The 250 ml round-bottom flask with a 24/40 ground edge, and glass beads, was previously dried in an oven at 105 ºC for 1 hour and weighed. The samples were placed in the round-bottom flask. The fat present in the cacao material was extracted in a Soxhlet extractor by siphoning petroleum ether (distillation range 30-60 °C; density at 15 °C 0.625-0.660 g ml^{-1}) for six hours. Petroleum ether was added in excess so that the fat extracted in the flask did not dry above the condenser. The fat was deposited at the bottom of the flask along with a small amount of petroleum ether and transferred to a beaker into a rotary vacuum evaporator to evaporate the petroleum ether. After evaporation, the fat (cream-coloured substance) that remained as a residue was weighed and expressed as a percentage.

Purine alkaloids

DCB milled (endosperm) was used for quantifying theobromine (THB) and caffeine (CAF) simultaneously with HPLC according to the method described by Brunetto et al. (2007).

Phenolic substances

The phenolic substances (-)-epicatechin (EPI) and (+)-catechin (CAT) contents of DCB (milled endosperm) were determined using the procedure described by Elwers et al. (2009). The total phenolic substances (PHE) of dried and milled cacao endosperm were determined using the Folin-Ciocalteu procedure described by Singleton and Rossi (1965) and Elwers et al. (2009).

Data analysis

The statistical procedures used in this study were performed using the R software (R Development Core Team 2021) and were described integrally by Loureiro et al. (2023).

Results and Discussion

According to the percentage of defects identified by the cut test of the cacao commercial classification (MAPA, 2008), all the observations in the sample correspond to DCB of Type I: mouldy (0-4%), smoky (0-1%), damaged by insects (0-4%), slate (0-5%), germinated (0-5%) and flattened (0-5%). Also, the moisture of the DCB is below the value of 8% established for Type I. The commercial classification of DCB by the cut test as Type I indicates that the post-harvest processing of fermentation and drying are within the standards required by Brazil and the world. The percents of brown, purple, and partially brown cacao beans corroborate this result (Table 1; Figure 1). None of these attributes were considered 'defects' when determining the commercial class of a batch of DCB. However, the presence of purple cacao beans (PCB) or bright PCB from the Forastero group or Trinitarian hybrids indicated either immature harvested fruit or insufficient fermentation (MAPA, 2008; Millena et al., 2023; Rahardjo et al., 2022). Brown cacao beans (BCB) indicated proper fermentation before drying (MAPA, 2008; Millena et al., 2023; Rahardjo et al., 2022). According to Hernández-Hernández et al. (2016), the dispersion of polyphenols during fermentation, followed by oxidation and reduction reactions due to external chemical substrates, produces this brown colour.

According to the Shapiro-Wilks test, the data obtained from the samples for the variables pH, LA, CAF, CAT, and EPI of DCB do not exhibit normality (Table 1). Loureiro et al. (2023) attribute these deviations from normality primarily to the reduction in sample observations. The data set used in this analysis corresponds to the sample observations from the study by Loureiro et al. (2023). Outliers in FAA, FI, and extreme outliers related to analytical errors were the main criteria used to eliminate 10 observations from the initial sample of 36 observations. Despite detecting some observations outside the interquartile range, some

Table 1. Sample analysis of biochemical attributes of cacao beans of the Porto Híbrido 16 (PH-16) clone cropping in the Humid region of southeastern Bahia, Brazil.

Type	Variable (Abbreviation or acronym, unit)	S-W test[1]		Outliers' detection	Descriptive statistics				Indexes[2]		Scatterplot analysis					
											FI quality range (ellipse cluster)					
		W	*p*-value	(boxplot)	Minimum	Maximum	Average	Standard deviation	GVF[3]	TAI[4]	*x* variable	*x* limits		*y* variable	*y* limits	
												x min	*x* max		*y* min	*y* max
NA	Potential of hydrogen (pH, dimensionless)	0.86	<0.01	3	5.82	6.57	6.05	0.19	0.94	0.72	NA	NA	NA	NA	NA	NA
CCA[5]	Brown cacao beans (BCB, %)	0.98	0.92	0	10.00	84.00	47.12	19.12	0.91	0.66	pH	5.79	6.05	BCB	8.63	94.98
	Purple cacao beans (PCB, %)	0.93	0.09	1	0.00	30.00	11.35	7.17	0.96	0.76	pH	5.78	6.05	PCB	-0.30 (zero)	18.68
	Partially brown cacao beans beans (PBCB, %)	0.97	0.64	0	5.00	90.00	40.42	22.03	0.93	0.71	pH	5.78	6.05	PBCB	-10.36 (zero)	85.60
OAC[6]	Acetic acid (AA, g kg^{-1})	0.99	1.00	0	9.27	33.84	22.14	5.92	0.93	0.73	pH	5.78	6.05	AA	3.49	41.42
	Lactic acid (LA, g kg^{-1})	0.88	<0.01	2	4.84	18.45	9.09	2.94	0.95	0.74	pH	5.81	6.03	LA	2.35	13.59
SCA[7]	Sucrose (SUC, g kg^{-1})	0.97	0.62	0	0.76	2.73	1.60	0.48	0.95	0.78	pH	5.79	6.05	SUC	0.50	1.96
	Fructose (FRU, g kg^{-1})	0.99	1.00	0	2.59	9.82	6.23	1.67	0.92	0.69	pH	5.79	6.05	FRU	0.54	9.32
	Glucose (GLC, g kg^{-1})	0.97	0.59	4	0.91	5.21	3.10	0.93	0.96	0.79	pH	5.79	6.04	GLC	-0.03 (zero)	5.52
SSU[8]	Total protein (PRO, g kg^{-1})	0.97	0.58	0	135.74	199.55	163.12	16.19	0.95	0.77	pH	5.79	6.05	PRO	120.40	210.05
	Total free amino acids (FAA, mg kg^{-1} FFDM[9])	0.97	0.61	0	10980.40	17276.53	13888.92	1754.92	0.94	0.76	pH	5.79	6.05	FAA	9224.22	19632.43
	Total lipids (LIP, g kg^{-1})	0.98	0.81	2	307.67	436.80	368.68	28.74	0.94	0.75	pH	5.79	6.05	LIP	315.03	430.89
PAL[10]	Theobromine (THB, g kg^{-1})	0.97	0.75	0	24.77	34.08	29.68	2.31	0.95	0.76	pH	5.78	6.06	THB	25.09	35.82
	Caffeine (CAF, g kg^{-1})	0.89	0.01	1	4.55	9.61	6.00	1.10	0.91	0.66	pH	5.80	6.04	CAF	3.93	6.83
PSU[11]	(+)-Catechin (CAT, g kg^{-1})	0.84	<0.01	1	0.65	5.72	1.89	1.16	0.95	0.76	pH	5.78	6.06	CAT	-0.52 (zero)	3.98
	(-)-Epicatechin (EPI, g kg^{-1})	0.86	<0.01	3	2.22	16.58	6.14	3.15	0.95	0.74	pH	5.78	6.05	EPI	2.52	8.32
	Total phenolic substances (PHE, g kg^{-1})	0.99	0.97	1	45.99	92.38	68.63	11.18	0.95	0.77	pH	5.78	6.05	PHE	52.55	84.19

[1]Shapiro-Wilks test at 5% significance level. [2]Indexes for assessing class intervals (Armstrong et al., 2003; Bivand et al., 2020). [3]Goodness of variance fit. [4]Tabular accuracy index. [5]Color characterization by the cut test of commercial classification. [6]Organic acids. [7]Simple carbohydrates. [8]Structural substances. [9]Fat free dry material. [10]Purine alkaloids. [11]Phenolic substances.

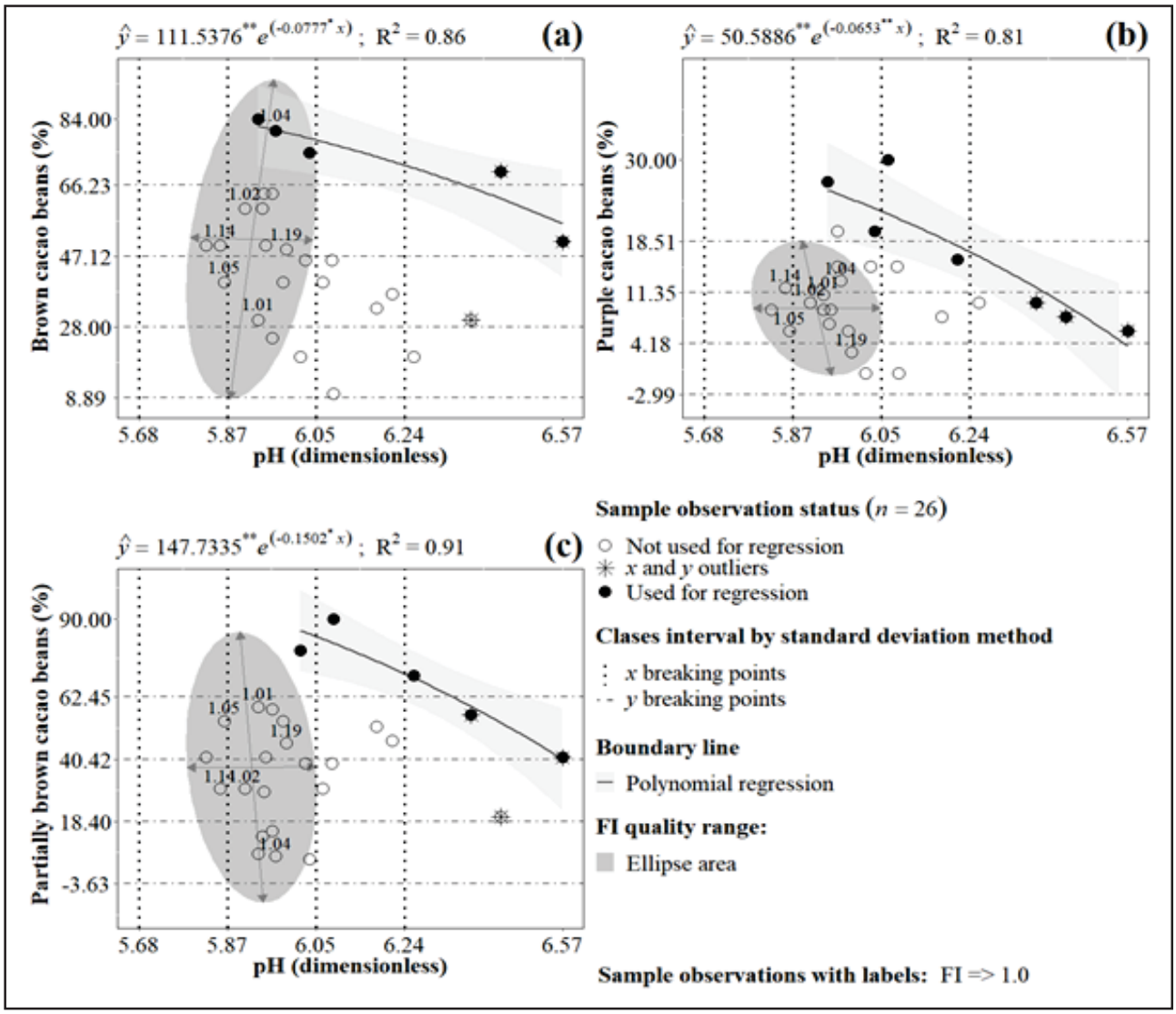

Figure 1. Boundary lines fitting of the polynomial regressions of the colour characterization by the cut test of commercial classification as a function of pH of dry cacao beans. Significance level of regression coefficients by the F test: $p \leq 0.01$ (**), and $0.01 > p \leq 0.05$ (*). Coefficient of determination (R^2). A confidence interval at 95% level. Ellipse area calculated with multivariate normal distribution. Porto Híbrido 16 cacao clone; Humid region of southeastern Bahia, Brazil.

were retained as they did not represent typical chemical analysis errors and did not affect the frontier regression trend (Table 1; Figures 1-4).

The internal variability of the sample observations classified into four intervals was robust based on the GVF and TAI indices (Armstrong et al., 2003). The average was used as a class limit on the x- and y-axis of Figures 1 to 4, as explained by Evans (1977). GVF values equal to or higher than 0.90 for all variables in the graphical analysis (Table 1) indicate a good fit for the method of determining class intervals (Smith, 1986). Only the variables BCB (Figure 1a), FRU (Figure 2d), and CAF (Figure 4b) showed TAI values lower than 0.70, suggesting some information loss due to sample size reduction or an inadequate number of classes to capture internal variability.

Loureiro et al. (2023) classified the sample observations with FI values equal to or higher than 1.0 (Ilangantileke et al., 1991; Tchouatcheu et al., 2019) as a category of the retrospective fermentation process. They grouped these samples with the multivariate elliptical analysis technique (Murdoch et

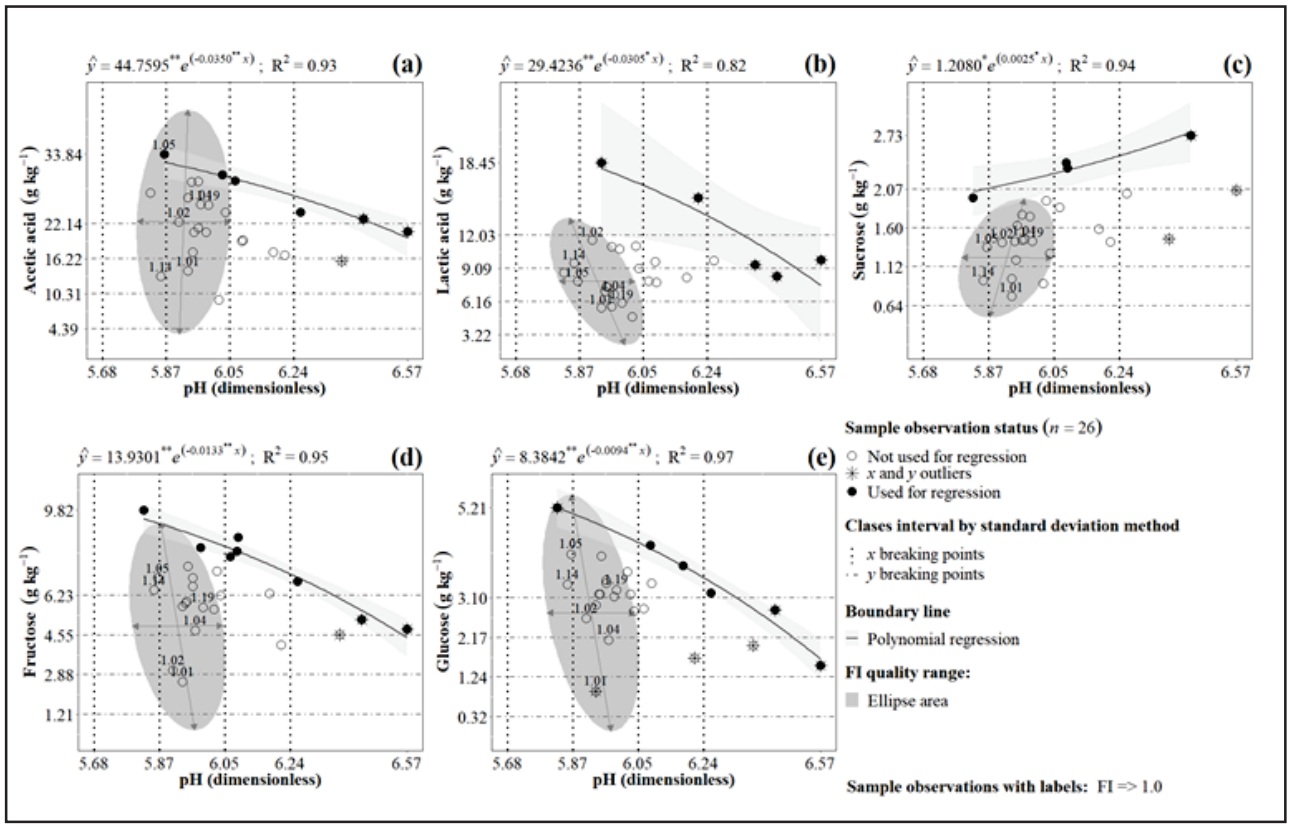

Figure 2. Boundary lines fitting of the polynomial regressions of the organic acids (a and b), and simple carbohydrates (c, d, and e) as a function of pH of dry cacao beans. Significance level of regression coefficients by the F test: $p \leq 0.01$ (**), and $0.01 > p \leq 0.05$ (*). Coefficient of determination (R^2). A confidence interval at 95% level. Ellipse area calculated with multivariate normal distribution. Porto Híbrido 16 cacao clone; Humid region of southeastern Bahia, Brazil.

al., 2022; Wickham et al., 2020), projecting the clusters on scatterplots of the relations between pH and FAA. This study also verified which pH quality ranges are related to FI values equal to or higher than 1.00 for the 16 analysed attributes (Table 1; Figures 1-4). The boundary lines estimated by exponential polynomial regressions show the distribution trend of the selected extreme values in each class interval (Figures 1-4). Except for sucrose (Figure 4a), all boundary regressions show an inverse relation between response variables and pH. The scatterplots (Figures 1-4) illustrate how the variables behave as a function of pH and FI (retrospective). The FI cluster in Figure 1b clearly shows that samples with a lower percentage of PCB are better fermented. Therefore, the PCB attribute is a potential indicator of fermentative quality when corroborated by pH analysis. Similar clustering patterns appear in the graphs for LA (Figure 2b) and SUC (Figure 2c). The ideal FI cluster in the AA samples suggests an ideal pH range associated with an optimal range for this variable (Figure 2b). Regarding SUC, it is evident that an increase in this attribute's concentration is not related to the good fermentative quality of the DCB.

Boundary line regressions are useful to represent biological processes or behaviours (Walworth et al., 1986; Webb, 1972, 1973). In this study, we use this technique to observe the biochemical processes in beans related to the monitoring of cacao quality. Furthermore, the elliptical cluster area (with the set of observations within the FI quality range) in the graphical analysis of bivariate relations, as well as the technique applied in multivariate analysis (Friendly et al., 2013; Jolicoeur, 1959), can be used to establish critical limits for these biochemical attributes for the same cacao genotype in different cropping sites.

The more density points observed in the scatterplot, the better the graphical analysis of the boundary line

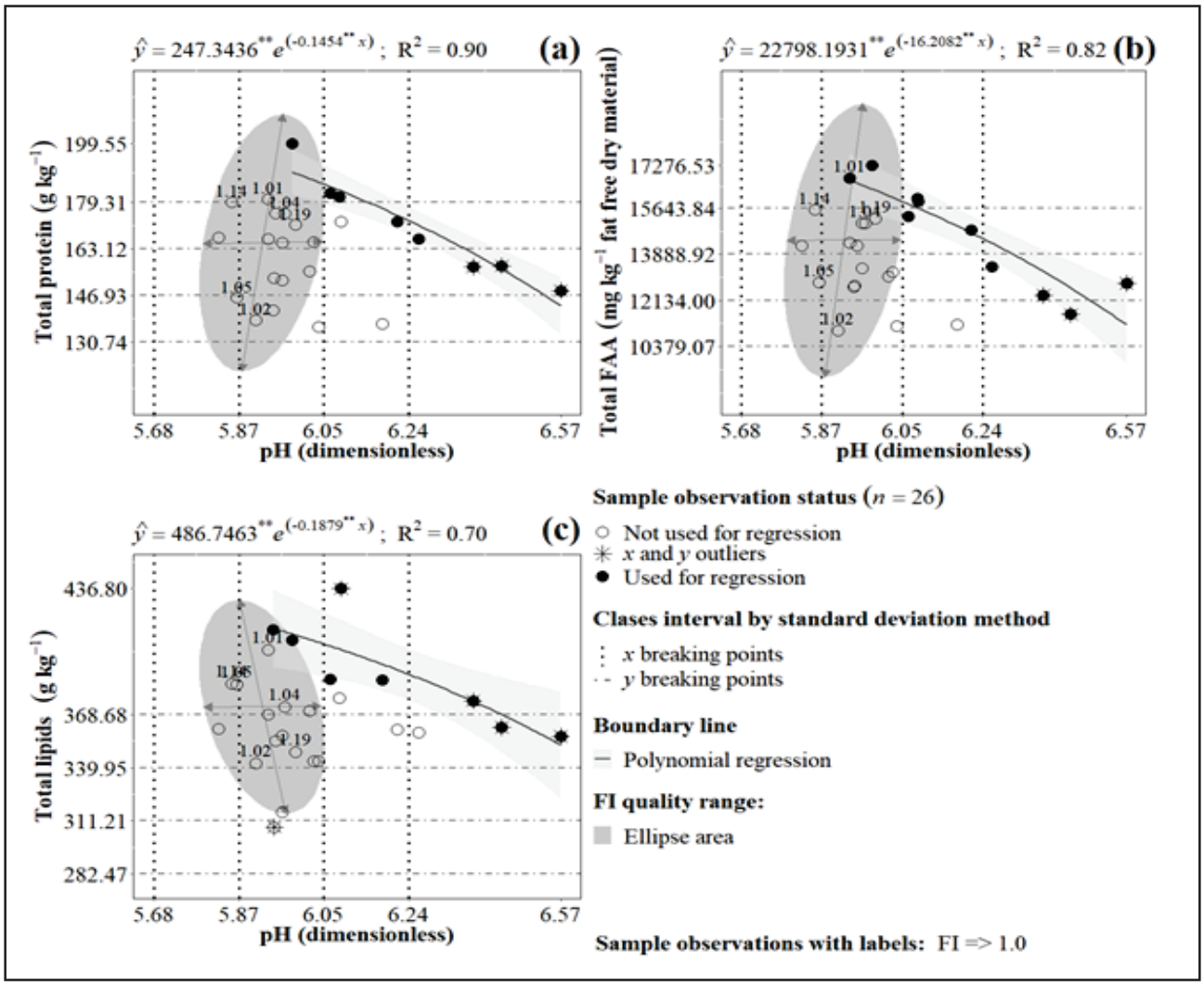

Figure 3. Boundary lines fitting of the polynomial regressions of the structural substances as a function of pH of dry cacao beans. Free amino acids (FAA). Significance level of regression coefficients by the F test: $p \leq 0.01$ (**). Coefficient of determination (R^2). A confidence interval at 95% level. Ellipse area calculated with multivariate normal distribution. Porto Híbrido 16 cacao clone; Humid region of southeastern Bahia, Brazil

will be (Walworth et al., 1986; Webb, 1972). In addition, large samples ($n > 100$) seem more convincing because they meet the normal distribution criterion (Guttman and Smith, 1969). For example, the literature suggests detecting and eliminating outliers before performing a graphical analysis of boundary lines (Shatar and McBratney, 2004; Walworth et al., 1986), similar to the assumptions of an analysis of variance. However, an outlier in the interquartile range should not always be eliminated and may sometimes be overlooked, especially when dealing with small samples ($n < 100$). When the sample is small, as in this study, the number of gaps increases, meaning there is not always an ideal density of observations. Therefore, each 'outlier' must be analysed in its analytical context to rule out the possibility that the values found are not commonly cited in the literature. Furthermore, increasing the number of repetitions of an experiment can result in unnecessary financial expense. For this reason, this study sought to solve the problem of critical limits through graphical techniques that are simple to execute and statistically feasible. Instead of eliminating all outliers from the sample and reducing the data set, they were identified graphically (Figures 1-4). Figures 1 to 4 show that the outliers are within the expected variability for the cacao bean attributes.

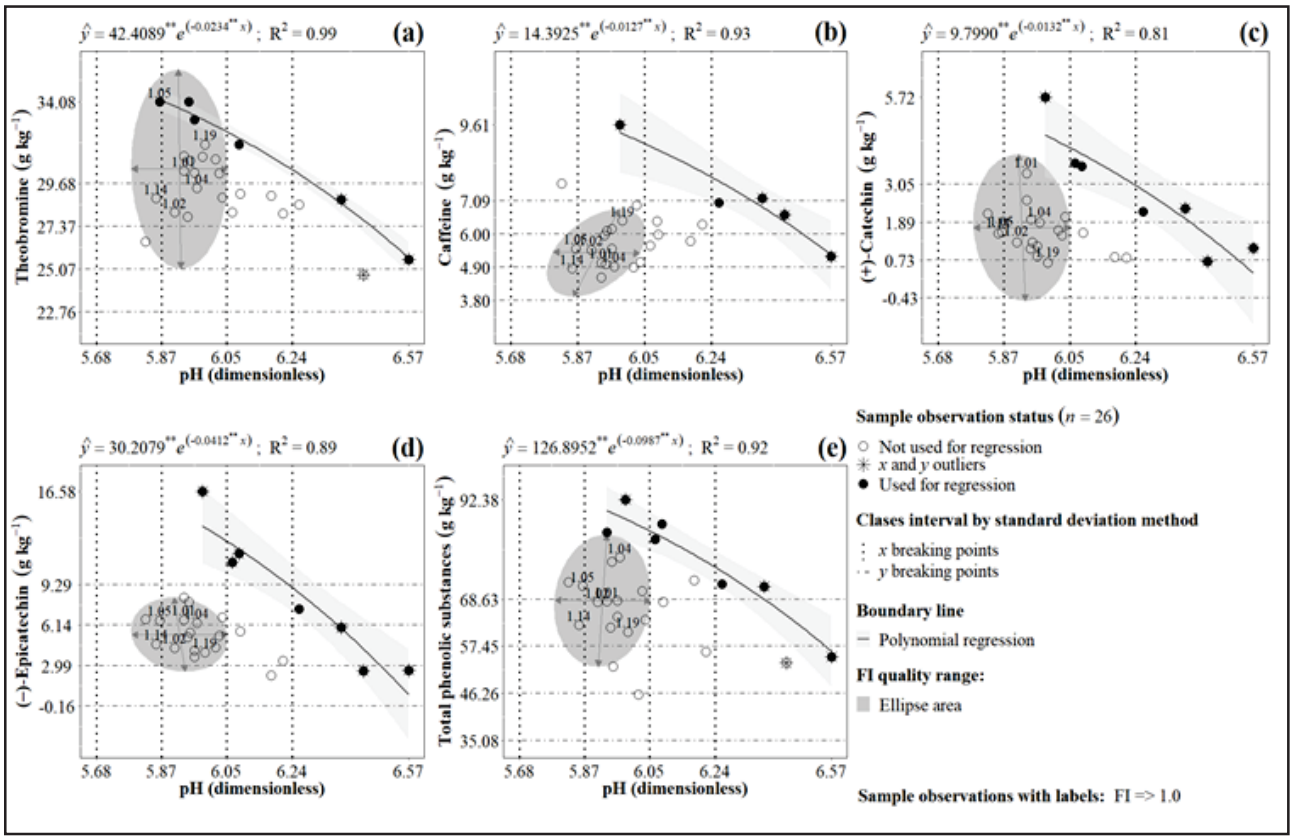

Figure 4. Boundary lines fitting of the polynomial regressions of the purine alkaloids (a and b), and phenolic substances (c, d, and e) with as a function of pH of dry cacao beans. Significance level of regression coefficients by the F test: $p \leq 0.01$ (**). Coefficient of determination (R^2). A confidence interval at 95% level. Ellipse area calculated with multivariate normal distribution. Porto Híbrido 16 cacao clone; Humid region of southeastern Bahia, Brazil.

On the other hand, the estimated FI quality range was able to simultaneously delimit the critical limits of pH and other biochemical attributes (Table 1; Figures 1-4). By averaging the pH range values of all DCB attributes (Table 1; Figures 1-4), we obtain the following critical limits 5.79 and 6.05. This range indicates a higher probability of best-fermented cacao samples, and, consequently, a greater potential for using the biochemical and sensory characteristics of the beans for the chocolate industry, for example. However, the FI quality range is two-dimensional, that is, pH is not the only limiting variable, and each response variable (*y*) has its critical limits (Table 1; Figures 1-7). In practice, the FI quality range is an ideal or optimal one, and from it, it is possible to establish other critical ranges such as "over fermented " or "under fermented". For example, the optimal range for total acidity lies between the values of 11.56 and 17.30 meq NaOH 100g^{-1}, and the range of lower acidity (< 11.56 meq NaOH 100g^{-1}) corresponds to an under fermented condition (Loureiro et al., 2023). The range of higher acidity (> 17.30 meq NaOH 100g^{-1}) may indicate over fermentation. But the extreme values estimated by the ellipse area may require an operational correction as is the case for PCB, PBCB, GLC, and CAT, whose new lower bounds will be zero (Table 1; Figures 1b; 1c; 2e; and 4c). Furthermore, the interpretation of extreme values must be carried out according to the biochemical behaviour of the variable in the fermentation and drying processes. SUC, for example, breaks down to form the other sugars FRU and GLC. Therefore, SUC values above the upper limit of the FI quality range (> 1.96 g kg^{-1}) may indicate under-fermentation. Thus, the interpretation of the upper limits for FRU and GLC carbohydrates is inverted; their upper limits within the FI quality range indicate over-fermentation.

As in this research, because all post-harvest processes (fermentation and drying) were controlled, the variability found in most of the attributes may be related to the plant nutrition, as well as the degree of maturity of the fruits, their storage time and transport, until the beginning of the fermentation process of unfermented cacao beans (Araujo et al., 2017; Cruz et al, 2013; Dang and Nguyen, 2019; Loureiro et al., 2017). PCB, LA, and SUC are variables whose limits of the FI quality range (*y*-axis) have a good discriminatory potential. That is, the values of the *y*-variable corroborate the values of the pH range. For the other variables, only the pH range allows for discriminating the fermentative quality of the sample observations. The interpretation of these critical limits must be based on the specialised literature; however, only new experimental treatments could induce variability capable of improving and correcting the FI quality range area as a two-dimensional probability space.

Conclusion

The projection of the elliptical clustering area in two-dimensional space, representing the response of biochemical attributes to variations in the pH of dry cacao beans, proved to be a viable technique for monitoring samples within the optimal fermentation index range (1-1.4), a variable arising from the end of the controlled fermentation process.

The pH values between 5.79 and 6.05 represent an overall range in which the biochemical attributes of dry cacao beans (organic acids, simple carbohydrates, structural substances, purine alkaloids, and phenolic substances) are found in an optimal condition of fermentation quality.

Boundary line regressions represent the response trends of biochemical attributes as a function of pH prediction, a variable sensitive to different fermentation and drying conditions of cacao beans. However, the classification of samples outside the two-dimensional quality range must be interpreted within the biochemical context of each variable to accurately represent conditions of over- and under-fermentation.

In addition to the pH range, the upper limits of the FI quality range for purple cacao beans, lactic acid, and sucrose attributes have good discriminatory potential. The critical limits of cacao quality were calibrated for the fermented and dry beans from the PH-16 clone; therefore, they should not be extrapolated to different cultivars or genetic varieties.

Acknowledgements

This paper is part of the project "Linking soil quality and cocoa quality in Bahia, Brazil" and, for running the fundamental steps of this research, the author Quintino R. de Araujo was supported by the Brazilian National Council for Scientific and Technological Development (CNPq) (grant Proc. 201739/2008-9 PDE) for performing Postdoctoral fellowship. The project SCA 58-1245-3-237F of USDA-ARS-FUNPAB partially supported research activities.

Literature Cited

APRIYANTO, M. et al. 2016. Study on effect of fermentation to the quality parameter of cocoa bean in Indonesia. Asian Journal of Dairy and Food Research. 35(2):160-163. https://doi.org/10.18805/ajdfr.v35i2.10724

ARAUJO, Q. R. et al. 2014. Cocoa Quality Index – A proposal. Food Control 46:49-54. https://doi.org/10.1016/j.foodcont.2014.05.003

ARAUJO, Q. R. de et al. 2017. Impact of soils and cropping systems on mineral composition of dry cacao beans. Journal of soil science and plant nutrition 17(2):410-428.

ARAUJO, Q. et al. 2018a. Soil quality index for cacao cropping systems. Archives of Agronomy and Soil Science 64(13):1892-1909.

ARAUJO, Q. R. et al. 2018b. Impact of soils and cropping systems on biochemical attributes of dry cacao beans. Agrotrópica 30(3): 175-194. https://doi.org/10.21757/0103-3816.2018v30n3p175-194

ARAUJO, Q. R. et al. 2019. Cacao quality index for cacao agroecosystems in Bahia, Brazil. International Journal of Food Properties 22(1):1799-1814. https://doi.org/10.1080/10942912.2019.1675691

ARAUJO, Q. R. et al. 2021. Effect of Different Edaphic Crop Conditions on the Free Amino Acid Profile of PH-16 Dry Cacao Beans. Agronomy 11(8):1637.

ARMSTRONG, M. P.; XIAO, N.; BENNETT, D. A. 2003. Using Genetic Algorithms to Create

Multicriteria Class Intervals for Choropleth Maps. Annals of the Association of American Geographers 93(3):595–623. https://doi.org/10.1111/1467-8306.9303005

ASSOCIATION OF OFFICIAL AGRICULTURAL CHEMISTS - AOAC. 2005a. Official Method 31.1.08. In: Horwitz, W.; Latimer Jr., G. W. Official Methods of Analysis of AOAC International. 18th ed. AOAC International, Gaithersburg, Maryland.

ASSOCIATION OF OFFICIAL AGRICULTURAL CHEMISTS - AOAC. 2005b. Official Method 963.15-1973. In: Horwitz, W.; Latimer Jr., G. W. Official Methods of Analysis of AOAC International. 18th ed. AOAC International, Gaithersburg, Maryland.

ASSOCIATION OF OFFICIAL AGRICULTURAL CHEMISTS - AOAC. 2005c. Official Method 970.21-1974. In: Horwitz, W.; Latimer Jr., G. W. Official Methods of Analysis of AOAC International. 18th ed. AOAC International, Gaithersburg, Maryland.

BIN SAID, M. 1989. Some methods to determine the degree of fermentation in cocoa beans. In: Che Man Y. B.; Abdul Karim, M. N. B; Asbi, B. A. (Eds.). Conference on Food Processing: Prelude to the 90's. Faculty of Food Science and Biotechnology, UPM; pp. 41-51.

BRUNETTO, M. del R. et al. 2007. Determination of theobromine, theophylline and caffeine in cocoa samples by a high-performance liquid chromatographic method with on-line sample cleanup in a switching-column system. Food Chemistry 100: 459-67.

CINAR, Z.Ö. et al. 2021. Cocoa and cocoa bean shells role in human health: An updated review. Journal of Food Composition and Analysis. 103:104115. https://doi.org/10.1016/j.jfca.2021.104115

CRUZ, J. F. M. et al. 2013. Assessment of the fermentative process from different cocoa cultivars produced in Southern Bahia, Brazil. African Journal of Biotechnology 12(33):5218–5225. https://doi.org/10.5897/AJB2013.12122

DANG, Y. K. T; NGUYEN, H. V. H. 2019. Effects of Maturity at Harvest and Fermentation Conditions on Bioactive Compounds of Cocoa Beans. Plant Foods for Human Nutrition 74: 54–60.

ELWERS, S. et al. 2009. Differences between the content of phenolic compounds in Criollo, Forastero and Trinitario cocoa seed (*Theobroma cacao* L.). European Food Research and Technology 229:937–948. https://doi.org/10.1007/s00217-009-1132-y

EVANS, I. S. 1977. The Selection of Class Intervals. Transactions of the Institute of British Geographers 2(1):98. https://doi.org/10.2307/622195

FRANZEN, M.; BORGERHOFF MULDER, M. 2007. Ecological, economic and social perspectives on cocoa production worldwide. Biodiversity and Conservation 16(13):3835–3849. https://doi.org/10.1007/s10531-007-9183-5

FRIENDLY, M.; MONETTE, G.; FOX, J. 2013. Elliptical Insights: Understanding Statistical Methods through Elliptical Geometry. Statistical Science 28(1):1–39. https://doi.org/10.1214/12-STS402

GOURIEVA, M. B.; TSEREVITINOV, O. B. 1979. Methods for evaluating the degree of fermentation of cocoa beans (Patent No. 646254). USSR Patent.

GUTTMAN, I.; SMITH, D. E. 1969. Investigation of Rules for Dealing With Outliers in Small Samples from the Normal Distribution: I: Estimation of the Mean. Technometrics 11(3):527–550. https://doi.org/10.1080/00401706.1969.10490710

HERNÁNDEZ HERNÁNDEZ et al. 2016. Evaluation of different fermentation processes for use by small cocoa growers in Mexico. Food Science & Nutrition 4(5):690–695. https://doi.org/10.1002/fsn3.333

ILANGANTILEK, S. G.; WAHYUNDI, T.; GRACIA BAILON, M. A. 1991. Assessment Methodology to Predict Quality of Cocoa Beans for Export. Journal of Food Quality 14:481–496. https://doi.org/10.1111/j.1745-4557.1991.tb00088.x

JOLICOEUR, P. 1959. Multivariate Geographical Variation in the Wolf *Canis lupus* L. Evolution 13(3): 283–299. https://doi.org/10.2307/2406106

KHAIRUL BARIAH, S. 2014. Impact of Fermentation Duration on the Quality of Malaysian Cocoa Beans Using Shallow Box. Asia-Pacific Journal of Science and Technology 19:74-80.

KIM, H.; KEENEY, P. G. 1983. Method of Analysis for (-)-Epicatechin in Cocoa Beans by High Performance Liquid Chromatography. Journal of Food Science 48(2): 548-551.

LOUREIRO, G. A. H. A. et al. 2016. Influence of environmental factors on cacao biometric attributes. Agrotrópica 28(2):123-140. https://doi.org/0.21757/0103-3816.2016v28n2p123-140

LOUREIRO, G. A. H. A. et al. 2017a. Cacao quality: Highlighting selected attributes. Food Reviews International 33(4):382–405. https://doi.org/10.1080/87559129.2016.1175011

LOUREIRO, G. A. H. A. et al. 2017b. Influencia de factores agroambientales sobre la calidad del clon de cacao (*Theobroma cacao* L.) PH-16 en la región cacaotera de Bahia, Brasil. Ecosistemas y Recursos Agropecuarios 4(12): 579-587. https://doi.org/10.19136/era.a4n12.1274

LOUREIRO, G. A. H. A. et al. 2023. Quality range for free amino acids of cacao beans. Agrotrópica 35(1):5–14. https://doi.org/10.21757/0103-3816.2023v35n1p5-14

MINISTÉRIO DA AGRICULTURA E PECUÁRIA - MAPA. 2008. Instrução Normativa No. 38, de 23 de junho de 2008. Diário Oficial da União, Seção 1(117), 8. Publicado em 24 de junho de 2008. https://sistemasweb.agricultura.gov.br/sislegis/action/detalhaAto.do?method=v250964455

MILLENA, C. G. et al. 2023. Effect of fermentation stages on the nutritional and mineral bioavailability of cacao beans (*Theobroma cacao* L.). Journal of Food Composition and Analysis 115:104886. https://doi.org/10.1016/j.jfca.2022.104886

MURDOCH, D.; CHOW, E. D.; CELAYETA, J. M. F. 2022. Package 'ellipse'. Functions for Drawing Ellipses and Ellipse-Like Confidence Regions. R Package Version 0.4.3. R: The R Project for Statistical Computing. https://cran.r-project.org/web/packages/ellipse/ellipse.pdf

PLOETZ, R. 2016. The Impact of Diseases on Cacao Production: A Global Overview. In: Bailey, B., Meinhardt, L. (Eds). Cacao Diseases. Springer International Publishing. p. 33–59.

R DEVELOPMENT CORE TEAM. 2021. R: A Language and Environment for Statistical Computing (4.1.2). R Foundation for Statistical Computing. https://www.r-project.org/

RAHARDJO, Y. P. et al. 2022. Impact of controlled fermentation on the volatile aroma of roasted cocoa. Brazilian Journal of Food Technology 25: e2020270. https://doi.org/10.1590/1981-6723.27020

ROMERO-CORTES T. et al. 2013. Relationship between fermentation index and other biochemical changes evaluated during the fermentation of Mexican cocoa (*Theobroma cacao*) beans. Journal of the Science of Food and Agriculture 93(10):2596–2604. https://doi.org/10.1002/jsfa.6088

SHAHRIR, S.; DIMICK, P. S. 1986. Qualitative and quantitative measurements of cocoa beans fermentation. In: Proceedings of the Cocoa Biotechnology Symposium. Pennsylvania: Pennsylvania State University. pp. 55-78.

SHATAR, T. M.; MCBRATNEY, A. B. 2004. Boundary-line analysis of field-scale yield response to soil properties. J. Agric. Sci. 142(5):553–560.

SMITH, R. M. 1986. Comparing traditional method for selecting class intervals on choropleth maps. The Professional Geographer 38(1): 62–67. https://doi.org/10.1111/j.0033-0124.1986.00062.x

SINGLETON, V. L.; ROSSI, J. A. 1965. Colorimetry of total phenolics with phosphomolybdic-phosphotungstic acid reagents. American Journal of Enology and Viticulture. 16:144–158. https://doi.org/10.5344/ajev.1965.16.3.144

SCHWAN, R. F.; SOUZA, S. M. M. 1986. Quantificação por HPLC de álcool etílico, ácido lático e ácido acético, produzidos na fermentação tradicional de cacau. Ilhéus, Bahia, CEPEC/CEPLAC.

TCHOUATCHEU, G. A. N. et al. 2019. Effect of cacao bean quality grade on cacao quality evaluation by cut test and correlations with free amino acids and polyphenols profiles. Journal of Food Science and Technology 56(5): 2621-2627. https://doi.org/10.1007/s13197-019-03749-y

WALWORTH, J. L.; LETZSCH, W. S.; SUMNER, M. E. 1986. Use of Boundary Lines in Establishing Diagnostic Norms. Soil Science Society of America Journal 50(1):123-128.https://doi.org/10.2136/sssaj1986.03615995005000010024x

WEBB, R. A. 1972. Use of the Boundary Line in the analysis of biological data. Journal of Horticultural Science 47(3):309-319.https://doi.org/10.1080/00221589.1972.11514472

WEBB, R. A. 1973. A possible influence of pedicel dimensions on fruit size and yield in strawberry. Scientia Horticulturae 1(4):321–330. https://doi.org/10.1016/0304-4238(73)90016-2

WICKHAM, H. et al. 2020. Package 'ggplot2'. Create Elegant Data Visualisations Using the Grammar of Graphics. R Package Version 3.3.3. R: The R Project for Statistical Computing. https://cran.rproject.org/web/packages/ggplot2/ggplot2.pdf.