

**Contemporânea**

*Contemporary Journal*

3(3): 2202-2227, 2023

ISSN: 2447-0961

Artigo

## **ESTUDO DA DIFERENCIAÇÃO DAS ESPÉCIES DE ANDIROBA: *CARAPA GUIANENSIS* E *CARAPA PROCERA* ATRAVÉS DO MÉTODO QUIMIOMÉTRICO**

STUDY OF THE DIFFERENTIATION OF ANDIROBA SPECIES: *CARAPA GUIANENSIS* AND *CARAPA PROCERA* THROUGH THE CHEMIOMETRIC METHOD

DOI: 10.56083/RCV3N3-057

Recebimento do original: 17/02/2023

Aceitação para publicação: 13/03/2023

### **Cláudia Cândida Silva**

Doutora em Físico-Química

Instituição: Universidade do Estado do Amazonas (UEA)

Endereço: Av. Darcy Vargas, 1200, Parque Dez de Novembro, CEP: 69050-020, Manaus - AM

E-mail: claudiacsbr@gmail.com

### **Rayssa Chripim de Souza**

Mestranda em Clima e Ambiente

Instituição: Universidade do Estado do Amazonas (UEA)

Endereço: Av. Darcy Vargas, 1200, Parque Dez de Novembro, CEP: 69050-020, Manaus - AM

E-mail: rchripim97@gmail.com

### **João Baptista da Silva Ferraz**

Doutor em Ciências Florestais

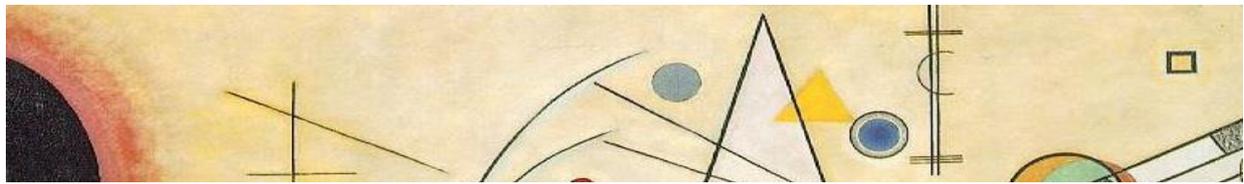
Instituição: Instituto de Pesquisa da Amazônia (INPA)

Endereço: Av. André Araújo, 2936, Petrópolis, CEP: 69067-375, Manaus - AM

E-mail: jferraz@inpa.gov.br

**RESUMO:** O presente trabalho teve como justificativa principal o estudo da diferenciação das espécies *Carapa guianensis* e *Carapa procera* com a utilização de dados de fluorescência de raios-X por dispersão de ondas. Estas são comumente chamadas de andiroba e andirobinha, respectivamente. Possuem distinções dificilmente perceptíveis a olho nú. O estudo da sua composição inorgânica é essencial no conhecimento mais aprofundado das espécies, uma vez que estas são cada vez mais utilizadas pela indústria. Avaliar as sementes quando cultivadas em solos diferentes é importante,

2202



uma vez que é necessário considerar fatores geológicos, climáticos e temporais na sua composição química como um todo. Levando em consideração as intensidades de metais detectados e ruídos gerados na construção do espectro da composição inorgânica, o objetivo do trabalho foi realizar a diferenciação, com o intuito inicial de separar as espécies através dos métodos quimiométricos, auxiliando na compreensão do sistema químico. Deste modo, comparou-se dados dos espectros das composições inorgânicas das sementes das duas classes de plantas, em solos distintos. Obteve-se como resultado que as elas possuem alto grau de similaridade em suas composições elementares, agrupando-se amostras de espécies diferentes, bem como amostras cultivadas em solos distintos, dados estes comprovados em estudos quimiométricos tanto de Análise de Componentes Principais (PCA) quanto de Análise de Agrupamentos Hierárquico (HCA). Dessa forma, as espécies não se separam pelos motivos levantados como hipótese neste trabalho. A diferença entre elas parece estar relacionada a outros fatores, tais como composição orgânica e/ou a fisiologia de estruturação do vegetal.

**PALAVRAS-CHAVE:** Andiroba, Quimiometria, Composição Inorgânica.

**ABSTRACT:** The present study had as main justification the study of the differentiation of the species *Carapa guianensis* and *Carapa procera*, with the use of X-ray fluorescence data by wave dispersion. These are commonly called andiroba and andirobinha, respectively. They have distinctions hardly discernible to the naked eye. The study of their inorganic composition is essential in the more detailed knowledge of the species, since these are increasingly used by the industry. Evaluating the seeds when grown on different soils is important, since it is necessary to consider geological, climatic and temporal factors in their chemical composition as a whole. Taking into account the intensities of detected metals and noise generated in the construction of the spectrum of the inorganic composition, the objective of the work was to carry out the differentiation, with the initial start of separating the species through the chemometric methods, aiding in the understanding of the chemical system. Thus, data from the spectra of the inorganic compositions of the seeds of the two classes of plants were compared in different soils. As a result, they have a high degree of similarity in their elemental compositions, grouping samples of different species, as well as samples grown in different soils, as evidenced in chemometric studies of both Principal Component Analysis (PCA) and Hierarchical Grouping Analysis (HCA). In this way, the species are not separated by the reasons raised as hypothesis in this work. The difference between them seems to be related to other factors, such as organic composition and / or plant structural physiology.



**KEYWORDS:** Andiroba, Chemometrics, Inorganic Composition.



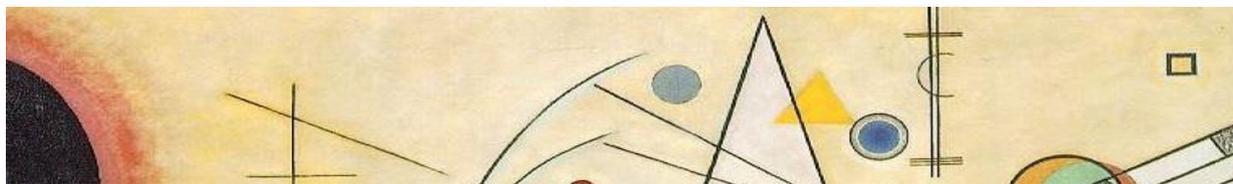
Artigo está licenciado sob forma de uma licença  
Creative Commons Atribuição 4.0 Internacional.

## 1. Introdução

A andiroba é um fruto conhecido por ter ampla distribuição geográfica, encontrada desde os neotrópicos até a África Tropical. Ocorre no Brasil e em países próximos, como Venezuela, Paraguai, Guiana Francesa, Peru, Suriname e Colômbia (PEREIRA, 2010; PINTO et al., 2010). Existem duas espécies atribuídas para a andiroba: *Carapa guianensis* Aubl. e *Carapa procera* D.C., tendo distinções dificilmente perceptíveis (FERRAZ et al., 2002). Ambas pertencentes à família botânica Meliaceae (QUEIROZ, 2007), que possui outras espécies como o cedro e o mogno (PINTO et al., 2010).

A andirobeira é uma árvore que possui o nome derivado do tupi-guarani ``*nhandi-rob*`` (*nhandi*: óleo e *rob*: amargo), palavra essa que relata o gosto amargo do óleo das suas sementes (CAVALCANTE, 2013). Também é uma árvore com excelente qualidade de madeira e de uso variável. (FERRAZ et al., 2002) As sementes de andiroba produzem um dos óleos mais aplicados na medicina popular amazônica (SHANLEY e MEDINA, 2005).

A espécie *C. guianensis* é utilizada por algumas populações juntamente com a espécie *C. procera*, no processo de obtenção do óleo de suas sementes. (MINISTÉRIO DA SAÚDE E ANVISA, 2015) A diferenciação entre as espécies *C. guianenses* e *C. procera* pode ser realizada através da observação do tamanho dos frutos e da mesma forma com relação às propriedades morfológicas (macroscópicas) na flor, semente e plântula (MINISTÉRIO DA SAÚDE E ANVISA, 2015). No entanto, apesar da extensa



literatura sobre estudo de andiroba, existem poucas referências que diferem as espécies *C. procera* e *C. guianensis*. Desse modo, ainda há muito a ser analisado sobre a distinção das duas, uma vez que são raramente trabalhadas separadamente e que podem auxiliar no manejo de comunidades que as utilizam. (FERRAZ et al., 2002)

Com o intuito de estudar a separação das espécies de andiroba, vários métodos analíticos podem ser empregados. A maioria deles gera uma infinidade de dados a serem devidamente processados (POPPI e SOUZA, 2012). Um estudo estatístico que pode ser utilizado é a quimiometria, que é um procedimento aplicado em resultados multivariados destinados a análises qualitativas (análise exploratória e reconhecimento de padrões) e quantitativas (calibração). (PARREIRA, 2003). Um dos métodos quimiométricos utilizados para identificar as semelhanças e as diferenças em vários tipos de amostras, agrupando-as e classificando-as, é o método “não supervisionado” que também é chamado de análise exploratória, que está baseado na validade das seguintes suposições: 1) As amostras do mesmo tipo são semelhantes; 2) Existem diferenças significativas entre diferentes tipos de amostras; 3) O conjunto de medidas disponíveis é capaz de detectar essas semelhanças e diferenças. Diante disso, dois métodos de análise exploratória são aplicados nos dados complexos, possuindo assim, gráficos de agrupamentos (FERREIRA, 2015). São eles: Análise de Componentes Principais (ACP) e Análise de Agrupamento Hierárquico (AAH).

Portanto, o problema científico deste trabalho refere-se ao uso sem diferenciação da andiroba. Dessa forma, como existe mais de uma espécie de andiroba sendo utilizada pela população, deve ser verificada a influência da composição inorgânica das espécies e a possível diferenciação das espécies, uma vez que elas são aplicadas em diversos setores produtivos da população.

Logo, como hipótese ao problema proposto tem-se que:



A espécie e o local de coleta são imprescindíveis para a interpretação da análise da composição, uma vez que pode influenciar no estudo presente.

Assim, o objetivo geral deste trabalho consiste em: tentar diferenciar as espécies *C. guianensis* e *C. procera* a partir das amostras de sementes em diferentes localidades. E tem-se os objetivos específicos: utilizar os resultados de fluorescência de raios-X por dispersão de ondas (FRXDO) obtidos em estudos anteriores de andiroba para a aplicação do método quimiométrico de estudo estatístico; analisar os dados obtidos no programa de Pirouette com a finalidade de gerar um modelo de agrupamento que diferencie as duas espécies de andiroba aqui estudadas; testar o modelo gerado com dados de amostras conhecidas para validação do modelo gerado; analisar qualitativamente o agrupamento gerado pelo software, baseado nos dados de fluorescência de raios-X por dispersão de ondas, procurando as variáveis que separam possíveis grupos de diferentes localidades de coleta.

## **2. Materiais e Métodos**

### **2.1 Amostras**

As amostras foram fornecidas pelo pesquisador João Baptista da Silva Ferraz, do Instituto Nacional de Pesquisas da Amazônia (INPA). Estas tiveram seus dados de composição química inorgânica obtidos em um projeto científico anterior, realizado dentro do nosso grupo de pesquisa. As devidas autorizações de coleta e manuseio das amostras foram apresentadas dentro daquele projeto.

Foram utilizadas 35 amostras da espécie *C. guianensis* e 10 amostras de *C. procera* em locais diferentes. Abaixo, na Tabela 1, encontram-se os pontos de coletas de cada espécie.

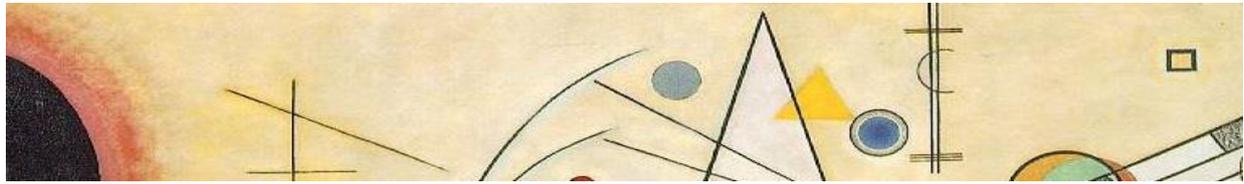


Tabela 1 – Pontos de coleta de cada espécie de andiroba.

Localidades	<i>C. guianensis</i>	<i>C. procera</i>
BR 174	x	
EMBRAPA		x
Estação Experimental de Silvicultura Tropical do INPA	x	x
Fazenda Mundico		x
Reserva Ducke	x	x

Estas amostras foram estudadas pelo método de fluorescência de raios – X por dispersão de onda em projeto de iniciação científica anterior, e estes dados foram utilizados para os estudos quimiométricos. Todos os cuidados de preparação das amostras e coleta de dados foram devidamente tomados, uma vez que, para estudos de comparação estatística, todos os dados devem estar exatamente nas mesmas condições. Foram testadas diferentes metodologias para a análise quimiométrica com o Pirouette, partindo-se dos dados totais de FRXDO.

## 2.2 Análise de Fluorescência de Raios – X

As sementes das duas espécies de andiroba foram recebidas com cascas secas embaladas em sacos plásticos.

Para analisar em fluorescência de raios – X, preparou-se a amostra da seguinte forma: para cada semente rompeu-se a casca e o descartou, os fragmentos menores da polpa da semente que estavam divididos foram triturados manualmente com auxílio de um almofariz e um pistilo de ágata até que se tornassem pó. Para cada trituração, os materiais utilizados eram higienizados com álcool etílico PA.

Através do auxílio de uma espátula plástica, pesou-se 4,0 g do ácido bórico com ajuda de uma balança analítica da Shimadzu, modelo AX200 e 1,0 g da amostra de cada semente. Primeiro introduziu-se o ácido bórico no porta-amostra para conduzir a prensagem e posteriormente a amostra da semente no mesmo. As amostras foram prensadas em uma prensa de



modelo PCA 30, da marca Teclago. Na prensagem, introduziu a amostra pesada no equipamento de forma centralizada e em seguida acrescentou uniformemente o ácido bórico por cima da amostra. Colocou-se a tampa e iniciou-se a prensagem. Posteriormente, realizou-se a análise desejada.

Utilizou-se o espectrômetro da Rigaku, modelo Supermini para a análise. O método consiste numa fonte de radiação de alta energia que promove a excitação dos átomos da amostra em análise. O átomo que estava no estado fundamental, tende absorver essa energia externa e favorece os elétrons a níveis mais energéticos. Na procura de estabilidade, o átomo excitado retorna seu estado fundamental naturalmente ocorrendo uma emissão de energia que está compreendida pela absorção, sendo uma propriedade específica de cada elemento químico que foi identificado e quantificado. (BECKHOFF et al., 2006)

A construção do espectrograma de fluorescência de raios-X se dá então com a detecção de cada fóton de energia liberado pela amostra em cada ângulo de giro do detector. Este gira em uma velocidade de  $0,01^\circ/\text{seg}$ . Varre-se um total de  $80^\circ$ . Diante disso, gera-se uma grande variedade de dados através dessa técnica, fornecendo assim, um ângulo que se relaciona com a intensidade de cada radiação característica, que é associada à quantidade do elemento químico no material ou ao ruído do efeito matriz. (BECKHOFF et al., 2006) Todos esses dados foram utilizados no estudo quimiométrico.

### 2.3 Método Quimiométrico

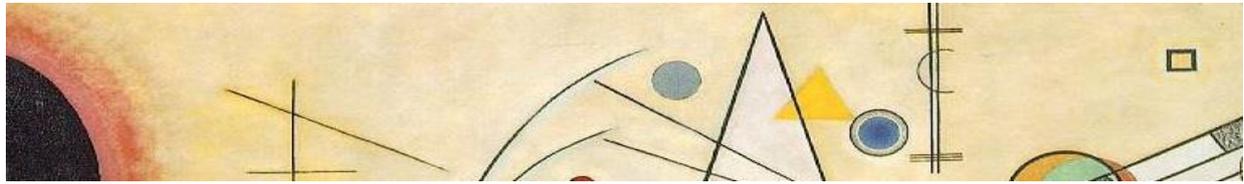
Através da obtenção dos resultados da análise multielementar, construiu-se uma tabela onde se relacionavam as amostras e as intensidades em cada ângulo de varredura. Um total de 4000 variáveis foram tabeladas. Este valor de dados, multiplicado pelo número de amostras (45 no total) justifica a utilização de um método matemático estatístico para avaliação precisa da diferenciação das espécies e locais de coleta.



O estudo prosseguiu utilizando o programa Pirouette com os seguintes seis passos:

- Examinar os dados: cada etapa foi possível verificar erros ou características importantes presentes nos dados. Neste, verificou-se a necessidade de algum pré-processamento dos dados;
- Pré-processamento dos dados: durante a análise de um problema foi verificada variações dentro do conjunto de dados que podem ocultar as informações mais relevantes para a análise. Assim, alguns métodos de pré-processamento são utilizados para reduzir fontes irrelevantes de variação nos dados. Estes são divididos em dois tipos: aplicados sobre as amostras, tais como normalização e correções de linha de base, ou aplicados sobre as variáveis em todos os pré-tratamentos.
- Estimar o modelo: após o pré-processamento dos dados, o próximo passo é obtenção do modelo quimiométrico utilizando programas computacionais apropriados;
- Examinar os resultados e validar o modelo: os modelos quimiométricos geram dados numéricos e gráficos e, com base nestes dados, é possível analisar e validar o modelo obtido. Nesta etapa foram utilizados conhecimentos matemáticos e químicos, pois algumas vezes o modelo obtido não é confiável e torna-se necessário a inclusão ou remoção de dados do sistema estudado;
- Fazer previsões com o modelo: será utilizado o modelo validado para prever o comportamento das amostras;
- Validar a previsão: a validação da previsão é muito importante, pois indicam desconhecidas que não se enquadram no modelo.

O programa disponibiliza 6 possibilidades de pré-processamentos. Através dos dados acima, usando todos os pré-processamentos para os passos seguintes, as análises de componentes principais e análise de agrupamento hierárquico foram levadas a cabo. As técnicas que agrupam

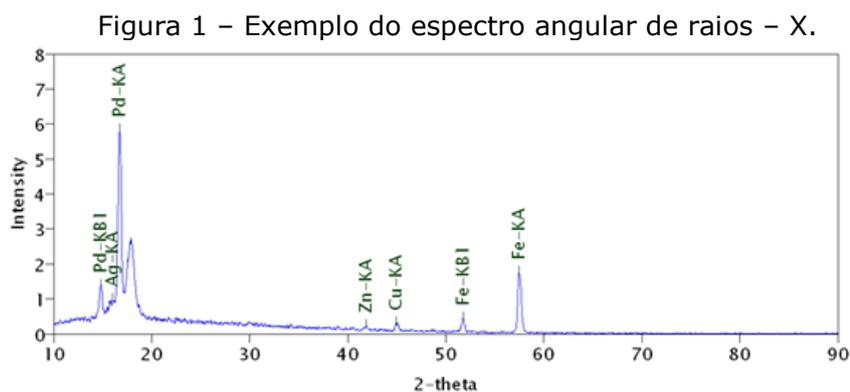


variáveis de forma estatística multivariada, tendo por objetivo reduzir a dimensão dos dados originais e permitindo a fácil visualização das informações mais importantes em um número menor de fatores, ou seja, as componentes principais e, a análise hierárquica de grupos, que consiste em calcular inicialmente as distâncias de cada amostra em relação às outras do conjunto de dados, visando seu agrupamento foram de extrema ajuda na avaliação do agrupamento ou não das amostras.

### 3. Resultados e Discussões

#### 3.1 Análise de Fluorescência de Raios – X

A análise por espectroscopia de fluorescência de raios – X é um método qualitativo e quantitativo baseado nas intensidades dos raios – X característicos emitidos pelos elementos que constituem a amostra. (DIAS et al., 2017) No intuito de caracterizar as amostras de sementes de andiroba, as amostras foram analisadas no equipamento de fluorescência de raios – X, originando-se espectros semelhantes a Figura 1 para cada amostra.



Com a utilização desta técnica é possível detectar os elementos majoritários, menores e traços presentes na amostra. (ERNST et al., 2012)



Uma vez que o espectro gerado é constituído não apenas do sinal referente aos elementos presentes, mas também da interação destes elementos entre si e com a matriz do material onde se encontram, conhecido como sinal-ruído. Sendo assim, o objetivo de utilizar também os sinais-ruídos é a verificação de uma relação significativa na possível relação entre as amostras das espécies dentro do estudo de quimiometria, uma vez que de acordo com Ferreira (2015), a medição de um sinal significa registro a magnitude da resposta de um instrumento em função de uma variável determinada pelo domínio da medida.

Logo os dados dos espectros como um todo, aplicado em técnicas matemáticas, possibilitam um estudo mais amplo de todas as variáveis geradas. A redução destas se dá de forma muito mais precisa, com a eliminação de dados indesejáveis de forma sistemática e reforçando as informações que realmente relevantes para o sistema em estudo. (FERREIRA, 2015)

Obtendo-se os espectros, utilizou-se os dados dos ângulos de cada intensidade para o estudo quimiométrico. Se as amostras são distinguíveis em qualquer agrupamento dos dados do espectro, pode-se concluir que as propriedades estão variando de forma significativa. Se as amostras são indistinguíveis no agrupamento dos dados, existe a possibilidade de terem propriedades não variando significativamente de acordo com os parâmetros aqui estudados.

### 3.2 Análise Exploratória dos Dados

Com o intuito de utilizar os dados do espectro como um todo, propõe-se os métodos quimiométricos não supervisionados para tentar separar as espécies, que são denominados análise exploratória, em razão que não se tem nenhuma informação estabelecida a respeito da classificação da amostra. Logo, foram agrupadas naturalmente com base na informação



contida nos dados experimentais, sem a preocupação de construir uma regra preditiva. (FERREIRA, 2015)

A parte que requer maior cuidado dentro de um estudo quimiométrico é a escolha dos possíveis parâmetros matemáticos a serem aplicados aos objetos de estudo. Utilizando o programa Pirouette, tem-se as possibilidades de escolher métodos de pré-processamento de dados, diferentes tipos de medidas de distância, métodos de ligação e ainda formas de escalonamento dos dados apresentados. Utilizou-se os dados em todos os pré-processamentos disponíveis no software para aplicação das amostras, a fim de buscar o melhor método, de forma experimental, para o agrupamento das amostras. Os resultados são específicos para cada conjunto de amostras. O pré-processamento é necessário porque vários algoritmos multivariados calculam os resultados buscando por padrões nas variáveis independentes aplicando métodos estatísticos matemáticos distintos. (INFOMETRIX, 2008) Para cada pré-processamento, gerou-se um gráfico com cada método de ligação e cada método matemático, para, assim, obter uma ilustração gráfica que explicita o resultado do cálculo em um formato de fácil interpretação. Para estudo de PCA, utilizou-se apenas os pré-tratamentos, sem alterações nos parâmetros de medida de distância, métodos de ligação e métodos de escalonamento. Para o HCA também foram testados métodos de ligação e de escalonamento distintos. A tabela 2 abaixo, demonstra os parâmetros utilizados na análise.

Tabela 2 – Parâmetros utilizados para quimiometria.

Algoritmo	HCA	PCA
Pré-processamento	Não pré-processado	Não pré-processado
	Autoescalamento	Autoescalamento
	Média Central	Média Central
	Escalonamento de alcance	Escalonamento de alcance
	Escalonamento de variância	Escalonamento de variância
	Pareto	Pareto



Distância Métrica	Euclideana	Não utilizado
Método de ligação	Simplex	Não utilizado
	Centro	
	Completo	
	Incremental	
	Mediana	
	Média do grupo	
	Flexível	
Métodos matemáticos de escalonamento	1º derivada	Não utilizado
	2º derivada	
	Alinhar	
	Linha de base	
	Dividir	
	Log10	
	MSC	
	Multiplicar	
	Normalizar	
	Smooth	
	SNV	
	Subtrair	

### 3.3 Análise de HCA

Para análises de HCA, resultaram-se gráficos em um esquema bidimensional, os dendogramas, que possuem uma escala de medida, que varia entre 0 a 1. Quanto mais próximas de 1, mais similar são as amostras agrupadas, quando observados seus agrupamentos.

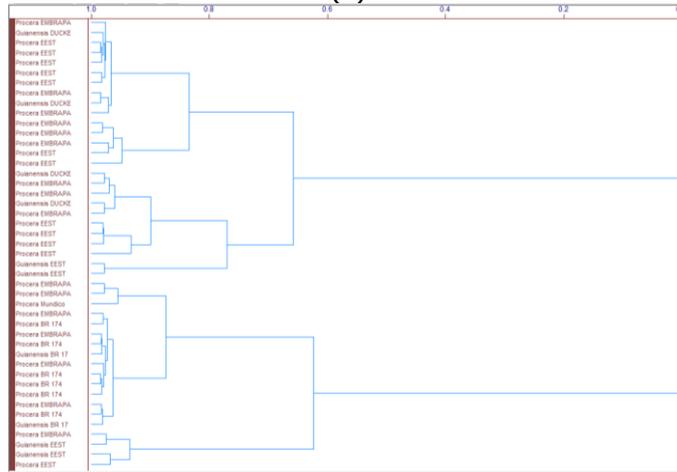
Nos gráficos de HCA, as distâncias entre pares das amostras são calculadas. Quando as distâncias entre as amostras são relativamente pequenas, isso implica na maior semelhança entre elas. As amostras que não são semelhantes estão separadas por distâncias relativamente grandes. (INFOMETRIX, 2008)

Com a observação dos gráficos gerados, verificou-se que cada pré-processamento modificava a construção dos gráficos com tamanho de ramos diferentes, porém, todos resultavam em algo comum. Dessa forma, o

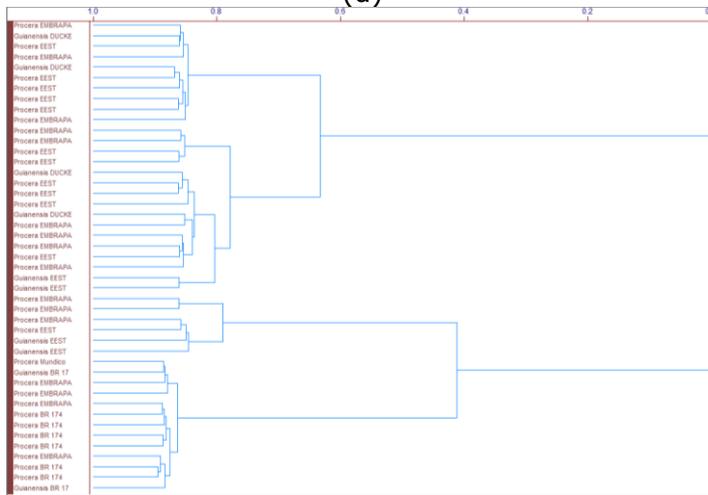




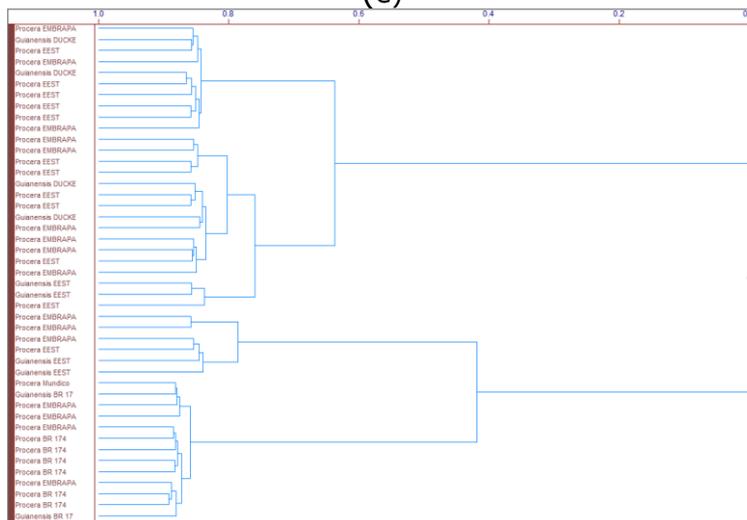
(c)



(d)

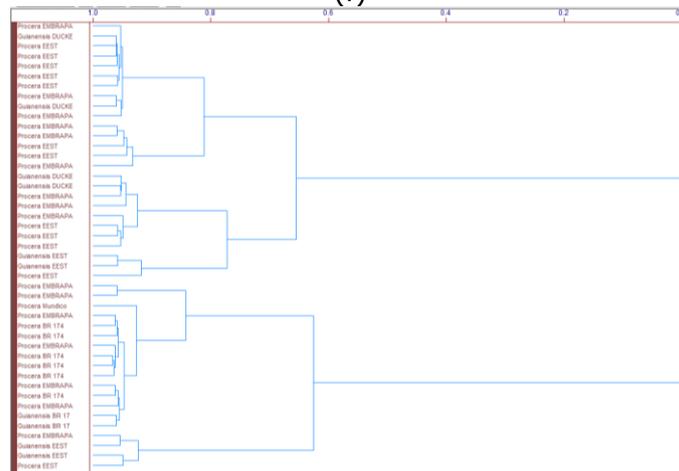


(e)





(f)



É possível observar a formação de dois grupos nos gráficos de cada pré-processamento, sendo que em cada grupo ocorre o agrupamento das amostras, ou seja, existe algo em comum que as une. Havia-se suposto que tal agrupamento se daria com cada conjunto contendo apenas amostras de uma dada espécie. Isso não foi verificado em nenhum dos grupos para nenhum dos métodos aplicados. Também se levantou a hipótese deste agrupamento ter se dado por diferentes tipos de solo, mas também não se trata deste parâmetro.

Observa-se que o não pré-processado teve o mesmo gráfico da mediana e o autoescalamento também apresentou congruência com de variância, sendo que ambas tiveram similaridades do resultado obtido em escalonamento de alcance, e o pareto teve o dendograma mais distante em relação aos outros. Porém, observa-se que todas as amostras de BR174 ficaram apenas em um grupo, o que pode explicitar uma tendência entre as espécies de absorverem elementos químicos de forma mais regular dependendo da composição do solo, mas não de forma suficientemente diferente para uma separação clara entre as amostras aqui estudadas. (EPSTEIN, 1975)

Os métodos matemáticos trabalhados não influenciaram na modificação dos gráficos, ou seja, não teve influência neste presente estudo



para cada pré-processamento realizado e cada método de ligação estabelecido.

Ocorrência esta, a qual já apareceu em alguns estudos cujo propósito era o mesmo, isto é, separação de dois grupos por meio de métodos quimiométricos. Exemplificando melhor esse fato, os estudos de separação entre dois tipos de arroz promovidos por Yun-Soo Yeo (et al., 2016) utilizaram de métodos de quantificação por meio de colunas de cromatografia preenchidas.

Ainda que possua uma quantificação diferente da que foi elaborada nesse estudo, obteve-se, portanto, que a quimiometria por meio do método não supervisionado de HCA mostrou que ambos os tipos de arroz não poderiam ser separados em agrupamentos visível e matematicamente distintos. Fator que ocorre também nessa análise, dado que, como já mencionado, houve um distanciamento relativamente pequeno em todos os pré-processamentos utilizados.

### 3.5 Análise de PCA

Dizemos que na projeção dos dados, que ocorre com a utilização do método não supervisionado de PCA, houve uma “compressão”, ou seja, uma perda controlada de informações que é compensada pela melhor compreensão do conjunto de dados. Esta é obtida fazendo combinações lineares das variáveis originais, de maneira a agrupar aquelas que fornecem informações semelhantes. Assim, tem-se como resultado, um novo conjunto de variáveis com propriedades mais representativas do conjunto e específicas. Essas novas variáveis são chamadas de componentes principais (do inglês *Principal Components* – PC), mas também são conhecidas por fatores. (FERREIRA, 2015)

Cada fator possui uma informação que não está contida em outro fator. A quantidade de informação dos dados originais que cada uma dessas

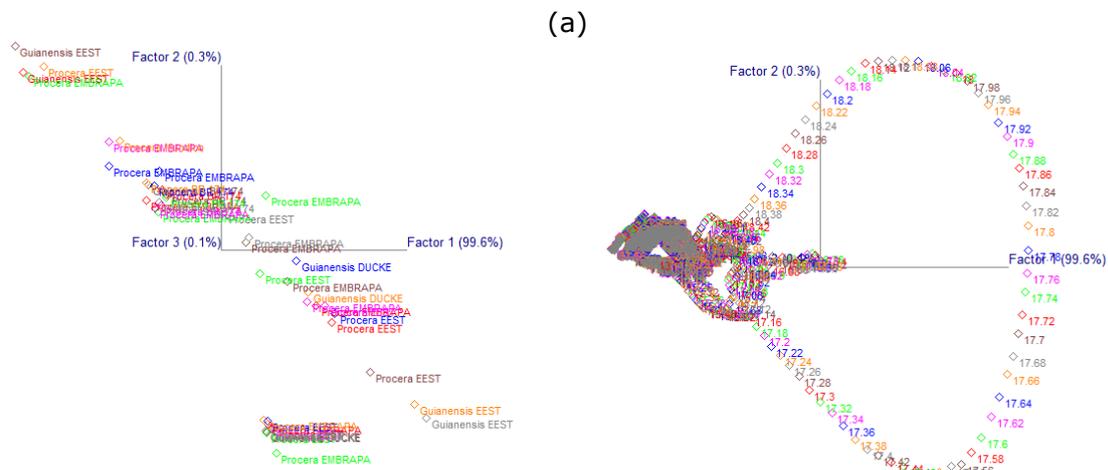


variáveis novas que descrevem são diferentes. O fator 1 é definido por descrever a máxima variância dos dados originais. O fator 2 possui a variância máxima dos dados no subespaço ortogonal ao fator 1 e os fatores posteriores são ortogonais às anteriores. (FERREIRA, 2015)

No software, a matriz dos dados é organizada de modo que as linhas sejam as amostras, também chamadas de *Scores*, e nas colunas as variáveis, ou *Loadings*. No caso dos espectros, nas linhas tem-se as amostras de sementes das espécies e nas colunas os valores das intensidades de cada ângulo da técnica analítica. (KUMAR e SHARMA, 2018)

Os gráficos obtidos de *score* e *loading* com os respectivos pré-processamentos, estão dispostos na Figura 3.

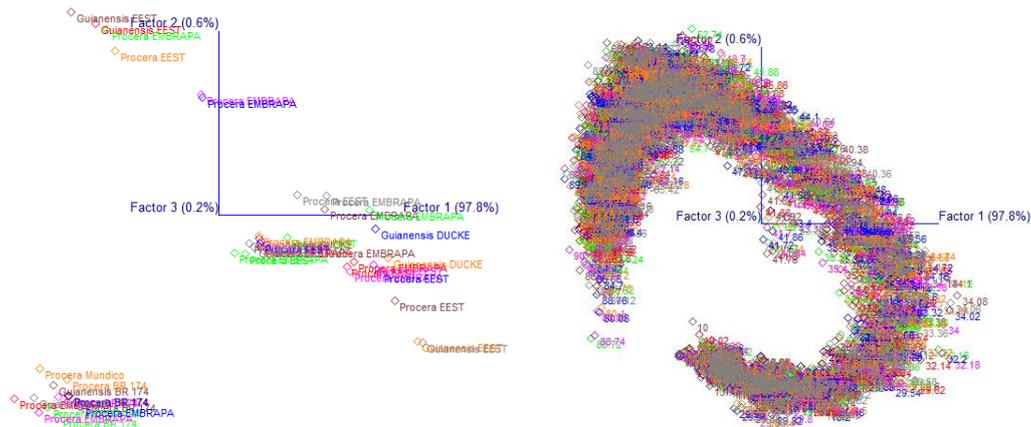
Figura 3 – Respectivamente os gráficos de *scores* e *loadings* de cada pré-processamento: a) nenhum, b) autoescalamento, c) mediana, d) escalonamento de alcance, e) escalonamento de variância e f) pareto



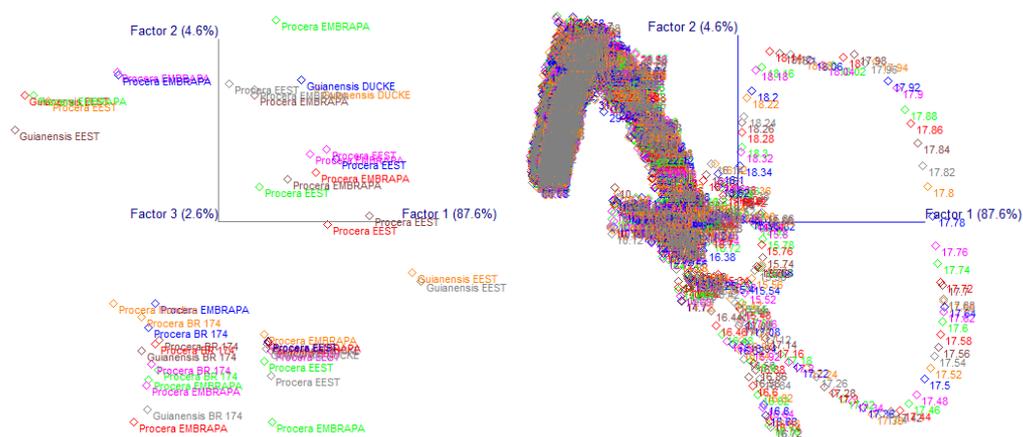




(e)



(f)



Os gráficos de *scores* são utilizados para comparar ou agrupar indivíduos em cada componente principal (VARELLA, 2011), ou seja, os *scores* correspondem as amostras nos fatores e os *loadings* a contribuição de cada variável. (PANERO et al., 2009)

O gráfico de *loading*, também é chamado de peso (VARELLA, 2011) deve estar acompanhado do gráfico de *score*, pois são eles que indicam por que as amostras foram agrupadas da maneira como estão dispostas no gráfico de *score*. (FERREIRA, 2015)

Dessa forma, no *score* de todos os pré-processamentos, teve-se a mistura das amostras através das similaridades entre elas, e nos fatores



obtidos é possível verificar que as espécies não se separaram, porque em um determinado grupo formado se tem as espécies *C. guianensis* e *C. procera* de solos distintos fazendo com que não haja percepção nítida de diferença estatisticamente entre as espécies ao que se refere a química inorgânica. E nos *loadings* as variáveis correspondiam os ângulos que possuíam suas intensidades, definindo assim, o que determinada a mistura de cada fator das amostras.

Pode ser observado também que em todos os tratamentos, os fatores deram diferentes apesar de nenhum terem separado as espécies. O gráfico não pré-processado possuiu maior variância no fator 1, da mesma forma a mediana, o escalonamento de alcance, escalonamento por variância e o pareto. O pré-processamento indicado é o autoescalonamento, uma vez que, teve-se o *score* com menor variância com 50,1% no fator 1, sabe-se assim, que quanto menor a variância mais próxima estará na média dos dados. Pode observar também, que neste gráfico relatado obteve-se três agrupamentos das amostras, onde todos haviam as amostras das duas espécies, percebe-se também que a espécie Procera EEST, coletada na Estação Experimental de Silvicultura Tropical do INPA se faz presente em dois agrupamentos, da mesma forma a Procera EMBRAPA se encontra nos três agrupamentos, indicando assim, que a localidade não influencia na sua composição.

Ao se comparar os gráficos de *loadings* e *scores* conjuntamente, pode-se enxergar alguns parâmetros que tendem a favorecer uma leve divisão entre as amostras nos gráficos de Pareto, todavia a razão desta não é dada pelas lógicas já analisadas, haja vista que a alça presente no gráfico, referente aos ângulos próximos à 17° apresenta a intensidade do Paládio (Rigaku Corporation, 2008), o qual é constituinte principal do tubo de emissor de raios-X do equipamento utilizado para as análises.

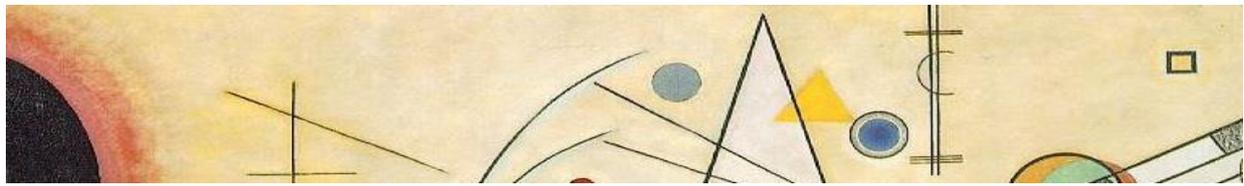
Desta forma, a variação na concentração de qualquer dos elementos químicos inorgânicos, matematicamente não tem relevância na separação dos conjuntos de amostras presentes nos resultados. Duas suposições



podem ser levantadas para a separação dos grupos: a densidade do material, uma vez que a região do gráfico que influencia é a reflexão da radiação do tubo, e a parte orgânica das amostras. Nenhuma destas possibilidades pode ser realmente estudada por métodos de fluorescência de raios-X. Assim trata-se apenas de uma conjectura para próximos estudos.

Este trabalho pode ser comparado com outro estudo cuja extração de dados foi feita a partir do obtido da Cromatografia Líquida de Alta Performance (HPLC). Segundo Liu W. (et al., 2016), é possível utilizar os métodos de PCA para validar os agrupamentos, tanto que o autor utiliza tal análise para separar diferentes amostras coletadas em distintas regiões da China de uma mesma planta medicinal do gênero *Potentilla*. Da mesma forma, o autor tem uma separação em grupos, onde a hipótese era de separação por espécies, o que não ocorre de forma clara, levando à necessidade de outros métodos de estudo para uma melhor explicação do agrupamento.

Em outro, estudo feito por Xu et al. (2018), uma avaliação de segurança do arroz geneticamente modificado foi conduzida por comparação com o arroz convencional. Algumas propriedades foram analisadas para identificar as equivalências biológicas e impactos do meio ambiente. Realizou-se a análise de similaridade baseado nos sinais eletrônicos (PEN3) com os métodos de HCA e PCA. Como no trabalho aqui realizado, o estudo de Xu et al. não obteve uma separação estatística entre as diferentes classes de arroz. Este fato não foi encarado de forma negativa pelos pesquisadores, uma vez que mostra, de forma clara, que o arroz alterado, levando-se em consideração os parâmetros avaliados, não é estatisticamente diferente do convencional, podendo ser comercializado de igual forma. Este é o mesmo ponto de vista que deve ser observado neste trabalho.



#### 4. Conclusão

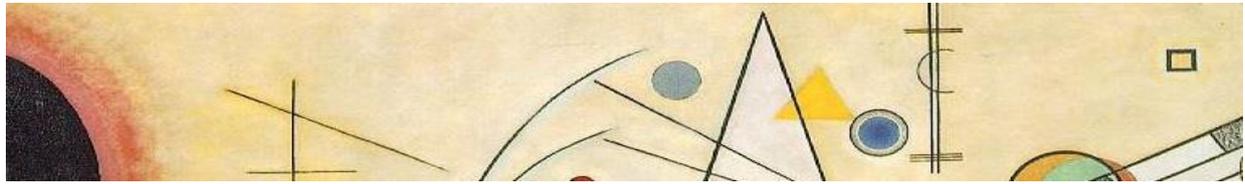
Assim, através da técnica analítica de fluorescência de raios - X, pode-se trabalhar com métodos quimiométricos, na qual geraram uma grande quantidade de dados, uma vez que se tratava do espectro como um todo.

Através do software Pirouette, manuseou-se os dados para uma geração de modelo com o intuito de verificação do agrupamento das espécies de *C. guianensis* e *C. procera* percebendo assim, que o método de ligação incremental e o tratamento mediana ilustra melhor o resultado obtido. Com os testes dos modelos gerados verificou-se que há a formação de grupos quimicamente distintos, uma vez que os gráficos construídos possuíam similaridades no agrupamento das amostras. Esta separação em grupos ocorre por fatores que não os levantados na hipótese deste trabalho.

A análise qualitativa dos agrupamentos, demonstrou que para o método HCA, o método de ligação constituído foi o de incremental para todos os pré-processamentos que resultaram gráficos parecidos, com os agrupamentos amostras e para o método PCA, demonstrou-se os agrupamentos das amostras em *scores* e em *loadings*, as variáveis que as grupavam.

No que se refere a aplicação prática deste trabalho, tem-se que as comunidades extrativistas, no que tange à composição química inorgânica, podem trabalhar com as duas espécies de *carapa* sem distinção, uma vez que a composição inorgânica não influencia no produto a ser comercializado por ela. Trata-se de vantagem importante, uma vez que isso facilita o trabalho deste grupo de pessoas, com uma etapa a menos (a de separação das espécies) na cadeia produtiva, sempre se levando em consideração apenas o fator da composição química inorgânica.

Por fim, o procedimento da monografia é satisfatório, uma vez que se verificou as similaridades das amostras em gráficos de HCA e PCA, onde não teve separação dos agrupamentos das espécies separadas, apenas das



amostras. Além de ser observado que através das localidades não ocorreu o agrupamento das espécies *C. guianensis* e *C. procera*, ou seja, as duas espécies de andiroba não se separam por meio dos métodos quimiométricos através do espectro da composição inorgânica.



## Referências

BECKHOFF, B.; KANNGIEBER, B.; LANGHOOF, N.; WEDELL, R.; WOLFF, H. **Handbook of Practical X-Ray Fluorescence Analysis**. 1 ed. Berlin-London: Springer, 2006.

CAVALCANTE, R. **Fitodontologia**. 1. ed. Rio Branco: Ed. do Autor, p. 268, 2013.

DIAS, B. N.; OLIVEIRA, D.; DOS ANJOS, M. J. A utilização e a relevância multidisciplinar da fluorescência de raios – X. São Paulo: **Revista Brasileira de Ensino de Física**, vol. 39, nº 4, 2017.

EPSTEIN, E. **Nutrição mineral de plantas: Princípios e perspectivas**. São Paulo: USP, 1975.

ERNST, T.; BERMAN, T.; BUSCAGLIA, J.; ECKERT-LUMSDON, T.; HANLON, C.; OLSSON, K.; PALENIK, C.; RYLAND, S.; TREJOS, T.; VALADEZ, M.; ALMIRALL, J. R. Signal to noise ratios in forensic glass analysis by micro X – rays fluorescence spectrometry. **X – ray spectrometry**: vol. 1, nº 43, p. 13 – 21, 2012.

FERRAZ, I.D.K; CAMARGO, J.L.C.; SAMPAIO, P.T.B. Sementes e Plântulas de andiroba (*Carapa guianensis* Aubl. e *Carapa procera* D.C.): Aspectos botânicos, ecológicos e tecnológicos. Instituto Nacional de Pesquisas da Amazônia. Manaus: **Acta Amazônica**, vol. 32, nº 4, p. 647 – 661, 2002.

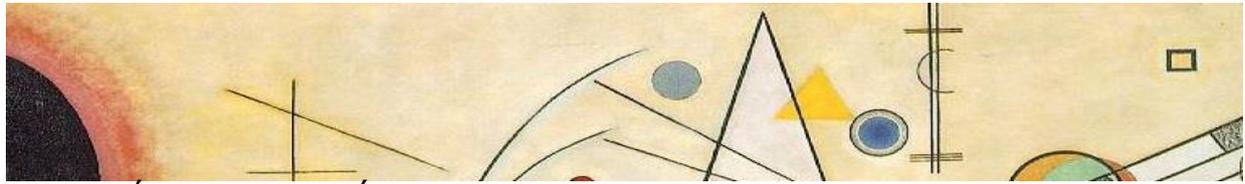
FERREIRA, M. M. C. **Quimiometria: Conceitos, Métodos e Aplicações**. Campinas – São Paulo: Unicamp, p. 496, 2015.

INFOMETRIX Inc. **Multivariate Data Analysis**. Version 4.0, 2008.

INFORMETRIX, Inc. **Pirouette Multivariate Data Analysis Software**. Version 4.5, 2014.

KUMAR, R.; SHARMA, V. Chemometrics in Forensic Science. **Trends in Analytical Chemistry**. Vol. 105, p. 191 – 201, 2018.

LIU, W.; WANG, D.; LIU, J.; LI, D.; YIN, D. **Quality Evaluation of *Potentilla fruticosa* L. by High Performance Liquid Chromatography Fingerprinting Associated with Chemometric Methods**. PLoS ONE, vol 11, nº 3, 2016.



MINISTÉRIO DA SAÚDE E ANVISA. **Monografia da espécie *Carapa guianensis* Aubl. (Andiroba)**. Brasília, 2015.

PARREIRA, T. F. **Utilização de métodos quimiométricos em dados de natureza multivariada**. Dissertação (mestrado) - Universidade Estadual de Campinas. São Paulo, 2003.

PANERO, F. S. ; VIEIRA, F. P.; CRUZ, Â. M.; MOURA, M. F. V.; DA SILVA, H. E. B. Aplicação da análise exploratória de dados na discriminação geográfica do quiabo do Rio Grande do Norte e Pernambuco. Universidade Federal do Rio Grande do Norte. São Paulo: **Eclética química**, vol. 34, nº 3, p. 33 – 40, 2009.

PEREIRA, M. R. N. **Fenologia e produção de sementes de andiroba (*Carapa guianensis*) Aubl. e suas implicações para o manejo**. Dissertação (Mestrado em Agronomia) - Universidade Federal de Roraima. Boa Vista, 2010.

PINTO, A.; AMARAL, P.; GAIA, C.; OLIVEIRA, W. **Boas práticas para manejo florestal e agroindustrial de produtos florestais não madeireiros: açaí, andiroba, babaçu, castanha-do-brasil, copaíba e unha-de-gato**. Belém, PA: Imazon, 2010.

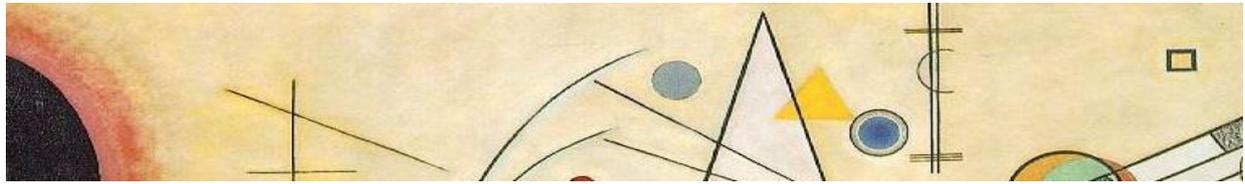
POPPI, R. J.; SOUZA, A. M. Experimento didático de quimiometria para análise exploratória de óleos vegetais comestíveis por espectroscopia no infravermelho médio e análise de componentes principais: um tutorial, parte I. São Paulo: **Química Nova**, vol. 35, nº 1, p. 223 –229, 2012.

QUEIROZ, J. A. L. **Guia prático de manejo florestal para produção de frutos de andiroba (*Carapa guianensis* Aubl.) e de outros produtos de valor econômico no estado do Amapá: 'A Floresta pode dar bons frutos'**. Macacá: IEPA, 2007.

SHANLEY, P.; MEDINA, G. **Frutíferas e plantas úteis na vida amazônica**. Belém: CIFOR, Imazon, 2005.

VARELLA, C. A. A. **Análise Multivariada Aplicada as Ciências Agrárias: Análise de Componentes Principais**. Universidade Federal Rural do Rio de Janeiro. Seropédica, 2011.

XU, S.; ZHOU, Z.; TIAN, L.; LU, H.; LUO, X. LAN, Y. Study of the similarity and recognition between volatiles of brown rice plant hoppers and rice stem based on the electronic nose. **Elsevier**, vol 152, p. 19-25, 2018.



YUN-SOO YEO, SEON-WOO OH, SOO-YUN PARK, SI-MYUNG LEE, SUNG-DUG OH, HYUN-SUK CHO, SOON-KI PARK, HYUN-JOO LEE, HEE-YUN KIM. Multivariate analysis for the safety assessment of genetically modified rices in the anti-nutrients and phenolic compounds. **International Journal of Food Science and Technology**, vol. 51, nº3, p. 765–776, 2016.