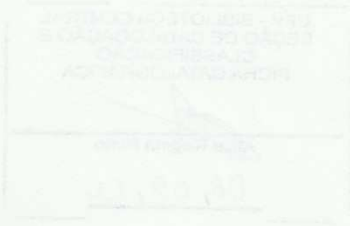


SOFRIMENTO FENIAS SAVANTO MATSIMBE

**UTILIZAÇÃO DE CARACTERÍSTICAS ÓPTICAS PARA ESTIMAR O  
TEOR DE ÓLEO E VOLUME DO MESOCARPO NOS FRUTOS DE  
MACAÚBA**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Fitotecnia, para obtenção do título de *Magister Scientiae*.

VIÇOSA  
MINAS GERAIS, BRASIL.  
2012



Ficha catalográfica preparada pela Seção de Catalogação e  
Classificação da Biblioteca Central da UFV

T

M442u  
2012

Matsimbe, Sofrimento Fenias Savanto, 1974-  
Utilização de características ópticas para estimar o teor  
de óleo e volume do mesocarpo nos frutos de macaúba /  
Sofrimento Fenias Savanto Matsimbe. – Viçosa, MG, 2012.  
xiv, 56f. : il. (algumas col.) ; 29cm.

Orientador: Sérgio Yoshimitsu Motoike.  
Dissertação (mestrado) - Universidade Federal de Viçosa.  
Inclui bibliografia.

1. *Acrocomia aculeata*. 2. Espectroscopia de  
infravermelho. 3. Processamento de imagens. 4. Macaúba.  
I. Universidade Federal de Viçosa. II. Título.


CDD 22. ed. 633.851

SOFRIMENTO FENIAS SAVANTO MATSIMBE


**UTILIZAÇÃO DE CARACTERÍSTICAS ÓPTICAS PARA ESTIMAR O  
TEOR DE ÓLEO E VOLUME DO MESOCARPO NOS FRUTOS DE  
MACAÚBA**

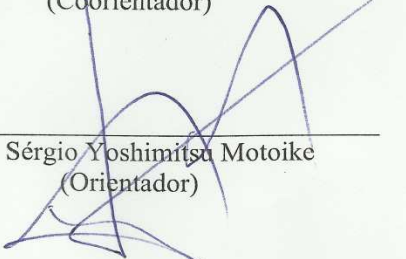
Dissertação apresentada a Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Fitotecnia, para obtenção do título de *Magister Scientiae*.

APROVADA: 27 de Agosto de 2012

  
Kacilda Naomi Kuki

  
Prof. Hélio Garcia Leite  
(Coorientador)

  
Prof. Francisco Assis de Carvalho Pinto  
(Coorientador)

  
Prof. Sérgio Yoshimitsu Motoike  
(Orientador)

*Aos*

*Meus pais (in Memorandum)*

*FENIAS SAVANTO MATSIMBE e REGINA ALICE MUIANGA*

*Ofereço*

*A*

*Minha esposa ISAURA ABÍLIO NOVELA e,*

*Meus filhos THANDIE MATSIMBE e NTIVE MATSIMBE*

*Dedico*

*Com*

*A minha família*

*MATSIMBES e NOVELAS*

*Compartilho.*

## AGRADECIMENTOS

Primeiro a DEUS pela vida e por iluminar todos os caminhos que trilho dia a dia, mês a mês, ano a ano a busca da minha formação e realização como Homem;

Aos governos de Moçambique e do Brasil pelos acordos assinados que me permitiram continuar com os estudos e, também, conhecer um povo maravilhoso e hospitaleiro, o Povo Brasileiro;

Ao CNPq pela concessão da bolsa de estudos;

Ao IIAM, minha entidade empregadora, pela autorização para continuar com os estudos, mas também pelo apoio moral e financeiro;

A Companhia do Vale pelo financiamento da passagem aérea de Moçambique para o Brasil;

A UFV, em especial o Programa de Pós-Graduação em Fitotecnia, pela aceitação e oportunidade de realizar este curso;

Ao meu Orientador Prof. Sérgio Yoshimitsu Motoike pela sábia orientação, dedicação, paciência, consideração, amizade e, sobretudo, por ser uma personalidade e pesquisador modelo para minha formação como pessoa e profissional;

Os meus Coorientadores Prof. Francisco Assis de C. Pinto, Prof. Hélio Garcia Leite e Prof. José Antonio Saraiva Grossi pela orientação, ensinamentos, dedicação, paciência, consideração, amizade e pelo exemplo que são para mim como personalidades e pesquisadores;

A Naomi Kuki e Carlos Nick pela amizade, atenção e contribuições no trabalho, mas, sobretudo pela aceitação para fazer parte da banca de defesa;

Ao amigo Gustavo Marcatti pelo precioso apoio na análise dos dados e desenvolvimento dos modelos, e também pela consideração, paciência e amizade;

Aos amigos Clenilso Sehnem Mota e Marley Lamounier pelo apoio, sobretudo nas fases iniciais do trabalho no campo, sugestões e ensinamentos;

Aos senhores Carlos e Cristiano por terem disponibilizado as suas fazendas (Acaiaca e Piranga) para os trabalhos de campo e pela consideração e amizade;

Ao Prof. José de Castro Silva pela amizade, apoio moral e material;

A Aurora Sato pela atenção, paciência e amizade;

Aos Professores dos Programas de Pós-Graduação em Fitotecnia, Engenharia Agrícola, Fisiologia Vegetal e Zootecnia pelas excelentes aulas;

Ao Grupo de pesquisa em Macaúba pela troca de experiência, interajuda, amizade e companheirismo, em especial a Ana Carla Brito, Angélica Plata, José Maria, Samuel Goulart e Leonardo Pimentel pelo apoio nas atividades do campo e no laboratório, Elma Souza pelo apoio material;

Aos funcionários do Setor da Fruticultura – UFV, especialmente Francisco, Telma, Beбето, Elimárcio, Edmar, Marcos e Claudinha pela pronta disponibilidade e ajuda, amizade e consideração;

Aos Colegas e estagiários dos Laboratórios de Pós-colheita - Setor da Fruticultura – DFT/UFV e do PROVISAGRO – DEA/UFV pela interajuda, paciência, consideração e amizade;

Aos colegas e amigos Anderson Costa, Wilker Medeiros e Eduardo Bottega pela amizade e apoio no experimento do processamento de imagens digitais;

Ao Prof. Herly Teixeira Dias e o seu grupo de pesquisa das florestas pela amizade e alegria e show de bola nos nossos amistosos de futebol;

Ao Sr. Marcelo Freitas das oficinas do DEA/UFV pelo equipamento por si idealizado e disponibilizado que foi bastante útil nas atividades práticas;

Aos grupos de estudantes Estrangeiros e Africanos em Viçosa pela troca de experiência, interação cultural e desportiva, amizade e companheirismo;

Aos meus companheiros do apartamento Prof. Ricardo Cruz e Atumane Nuro pelo ambiente saudável, amigo e familiar;

Ao meu grande Amigo, atleticano, Gilcemir pela amizade e atenção.

Aos meus pais Fenias Savanto Matsimbe e Regina Alice Muianga (*in Memorandum*) pela vida, carinho e sonhos e luta para ver os filhos formados;

A minha esposa Isaura Abílio Novela pelo amor e coragem para conciliar os seus estudos, trabalho e cuidar dos nossos filhos e do lar na minha ausência, os meus filhos Thandie e Ntive por tudo o que significam para mim;

As minhas duas famílias Matsimbe e Novela pelo carinho e apoio direto a minha pessoa e, também, de forma indireta aos meus filhos;

A minha imensa gratidão é extensiva a todos os demais que, pela traição da memória, não me lembrei de menciona-los, mas cuja contribuição quer de forma direta ou indireta foi relevante para o presente trabalho e para a minha formação.

*A TODOS UM ENORME KANIMAMBO.*

## **BIOGRAFIA**

SOFRIMENTO FENIAS SAVANTO MATSIMBE, filho de Fenias Savanto Matsimbe e de Regina Alice Muianga, nasceu em 15 de Abril de 1974, na cidade de Xai-Xai, província de Gaza, Moçambique.

Graduou-se como Bacharelato de Ciências Biológicas em Maio de 2007, pela Universidade Eduardo Mondlane na Cidade de Maputo, Moçambique.

Em Agosto de 2010, iniciou o Mestrado em Fitotecnia pela Universidade Federal de Viçosa e concluiu-o em Agosto de 2012.

*“Nos próximos 50 anos, cinco dos 10 maiores problemas terão que ser resolvidos pela agricultura, a saber: energia, água, alimentos, meio ambiente e pobreza”.*

***Alan Mac Diarmid** (Nova Zelândia)*

*Prêmio Nobel da Química 2000*



## SUMÁRIO

LISTA DE FIGURAS .....	X
LISTA DE TABELAS .....	XII
RESUMO .....	XIII
ABSTRACT .....	XIV
1 – INTRODUÇÃO GERAL .....	1
2 – REVISÃO DA LITERATURA.....	3
2.1 – Aspectos botânicos de macaúba.....	3
2.2 – Utilidades da macaúba .....	5
2.3 – Métodos analíticos usados para avaliação de produtos agrícolas .....	6
2.3.1 – Espectrometria do visível e infravermelho próximo .....	7
2.3.2 – Utilização da espectrometria do visível e infravermelho próximo em oleaginosas .....	9
2.3.3 – Processamento de imagens digitais e sua utilização na estimativa do volume de produtos agrícolas.....	11
3 – REFERÊNCIAS BIBLIOGRÁFICAS.....	12

### ARTIGO 1

---

#### **PREDIÇÃO DO TEOR DE ÓLEO DO MESOCARPO DA MACAÚBA USANDO ESPECTROMETRIA DO VISÍVEL E INFRAVERMELHO PRÓXIMO**

RESUMO .....	20
ABSTRACT.....	20
1 – INTRODUÇÃO .....	21
2 – MATERIAL E MÉTODOS .....	22
2.1 – Amostragem .....	22
2.2 – Mensuração dos dados espectrais.....	22
2.3 – Determinação dos valores de referência.....	23

2.4 – Ajuste do modelo VIS-NIR.....	23
2.5 – Análise estatística.....	24
3 – RESULTADOS E DISCUSSÃO.....	25
3.1 – Variabilidade nas amostras de macaúba.....	25
3.2 – Efeito do pré-tratamento dos dados espectrais.....	26
3.3 – Ajuste do modelo.....	27
3.3.1 – Determinação de variáveis latentes.....	27
3.3.2 - Calibração.....	28
3.3.3 – Validação externa.....	28
4 – REFERÊNCIAS BIBLIOGRÁFICAS.....	32

## **ARTIGO 2**

---

### **ESTIMATIVA DO VOLUME DO MESOCARPO DA MACAÚBA POR MEIO DO PROCESSAMENTO DE IMAGENS DIGITAIS**

RESUMO.....	36
ABSTRACT.....	36
1 – INTRODUÇÃO.....	37
2 – MATERIAL E MÉTODOS.....	38
2.1 – Amostragem.....	38
2.2 – Determinação dos valores de referência.....	38
2.3 – Estimativa do volume por meio do processamento de imagens digitais.....	39
2.3.1 – Processamento de imagens digitais.....	39
2.3.2 – Algoritmos do processamento de imagens digitais.....	40
2.4 – Análise estatística.....	42
3 – RESULTADOS E DISCUSSÃO.....	43
3.1 – Volume observado versus volume estimado.....	43
3.2 – Detecção de <i>outliers</i> .....	45

3.3 – Comparação dos algoritmos com o método padrão .....	47
3.3.1 – Comparação do MPID-esf com MDCA .....	47
3.3.2 – Comparação do MPID-elip com MDCA.....	49
3.4 – Teste de identidade dos algoritmos com o método padrão .....	52
4 – REFERÊNCIAS BIBLIOGRÁFICAS.....	53
CONCLUSÕES GERAIS .....	56

## LISTA DE FIGURAS

- Figura 1.** Características morfológicas da palmeira macaúba [*Acrocomia aculeata* (Jacq.) Lodd. ex Mart.]: a) vista geral da palmeira; b) inflorescência do tipo espádice; c) Cachos com frutos imaturos e maduros; d) vista longitudinal do fruto inteiro e após o corte transversal na qual são destacadas as frações específicas do fruto: epicarpo, mesocarpo, endocarpo e endosperma.....4
- Figura 2.** Radiação eletromagnética.....7
- Figura 3.** Esquema para desenvolvimento de um modelo quantitativo.....9

### ARTIGO 1

---

#### PREDIÇÃO DO TEOR DE ÓLEO DO MESOCARPO DA MACAÚBA USANDO ESPECTROMETRIA DO VISÍVEL E INFRAVERMELHO PRÓXIMO

- Figura 1.** Representação gráfica da determinação do número ótimo de VLs utilizado para ajustar o modelo, empregando RMSE (a) e  $R^2$  (b).....27
- Figura 2.** Correlação entre os valores estimados e observados pelos dois métodos: calibração (a) e validação externa (b).....29
- Figura 3.** Histogramas dos resíduos: calibração (a) e validação externa (b).....30

### ARTIGO 2

---

#### ESTIMATIVA DO VOLUME DO MESOCARPO DA MACAÚBA POR MEIO DE PROCESSAMENTO DE IMAGENS DIGITAIS

- Figura 1.** Processamento de imagens digitais: fruto de macaúba em RGB (a), escala de cinza (b) e binarizado (c).....40
- Figura 2.** Dimensões de uma esfera: D1 – diâmetro maior e D2 – diâmetro menor.....40
- Figura 3.** Calibração dimensional utilizando uma régua graduada e o *Software ImageJ* - determinação dos diâmetros em pixel e conversão para as medidas reais do fruto: quantidade de pixel por cm.....41
- Figura 4.** Dimensões de um elipsóide.....42
- Figura 5.** Representação gráfica *Box-and-Whisker*: detecção de *outliers* no conjunto de dados do MCDA (a), MPID-esf (b) e MPID-elip (c).....45
- Figura 6.** Valores do volume observados pelo MDCA *versus* estimados pelo MPID-esf.....48

- Figura 7.** Gráfico de *Bland-Altman* para a comparação do volume do mesocarpo dos frutos de macaúba determinado pelo MDCA e algoritmo MPID-esf, a linha tracejada na região central do gráfico indica a média das diferenças e as outras duas (regiões de cima e em baixo) indicam os limites de concordância a nível de 95%.....48
- Figura 8.** Valores do volume observados pelo MDCA *versus* estimados pelo MPID-elip.....50
- Figura 9.** Gráfico de *Bland-Altman* para a comparação do volume do mesocarpo dos frutos de macaúba determinado pelo MDCA e algoritmo MPID-elip, a linha tracejada na região central do gráfico indica a média das diferenças e as outras duas (regiões de cima e em baixo) indicam os limites de concordância a nível de 95% ....50

## LISTA DE TABELAS

### ARTIGO 1

---

#### **PREDIÇÃO DO TEOR DE ÓLEO DO MESOCARPO DA MACAÚBA USANDO ESPECTROMETRIA DO VISÍVEL E INFRAVERMELHO PRÓXIMO**

**Tabela 1.** Sumário da análise da variabilidade nas amostras de macaúba utilizadas nas séries de calibração e validação.....25

**Tabela 2.** Sumário do efeito do pré-tratamento sobre os dados espectrais.....26

**Tabela 3.** Sumário das estatísticas da calibração e validação externa do modelo de predição do teor de óleo do mesocarpo nos frutos de macaúba.....28

### ARTIGO 2

---

#### **ESTIMATIVA DO VOLUME DO MESOCARPO DA MACAÚBA POR MEIO DE PROCESSAMENTO DE IMAGENS DIGITAIS**

**Tabela 1.** Valores do volume (cm<sup>3</sup>) do mesocarpo dos frutos da macaúba determinados pelo MDCA, MPID-esf e MPID-elip, e os respectivos erros (e).....44

**Tabela 2.** Sumário dos resultados das estatísticas da comparação dos algoritmos com o método padrão.....47

**Tabela 3.** Sumário dos resultados do teste de identidade dos algoritmos com o método padrão .....52

## RESUMO

MATSIMBE, Sofrimento Fenias Savanto, M. Sc., Universidade Federal de Viçosa, Agosto de 2012. **Utilização de características ópticas para estimar o teor de óleo e volume do mesocarpo nos frutos de macaúba.** Orientador: Sérgio Yoshimitsu Motoike. Coorientadores: Francisco Assis de Carvalho Pinto, Hélio Garcia Leite e José Antonio Saraiva Grossi.

A macaúba [*Acrocomia aculeata* (Jacq.) Lodd. Ex Mart.] é uma oleaginosa rústica, de alta produtividade e múltiplas potencialidades, com grande demanda nas indústrias alimentícia, cosmética e, principalmente, a energética, em função da adequação para produção do biodiesel. Contudo, apesar do enorme potencial, a exploração da macaúba ainda resume-se ao extrativismo. Os estudos visando a sua domesticação e o conseqüente uso sustentável da espécie são insipientes, em função da falta de métodos práticos e eficientes que possam subsidiar os programas de melhoramento na seleção de genótipos de interesse. Assim, o objetivo deste trabalho foi desenvolver e propor métodos para estimar o teor de óleo e volume do mesocarpo no fruto da macaúba utilizando características ópticas. Para avaliação das metodologias foram conduzidos dois ensaios. No primeiro foram utilizadas 420 amostras para desenvolver um modelo de predição do teor de óleo do mesocarpo usando a espectrometria do visível e infravermelho próximo. Como referência foi usado o método *soxhlet*. No segundo, composto por 20 amostras, foram desenvolvidos algoritmos do processamento de imagens digitais para estimar o volume do mesocarpo, e como referência empregou-se o método do deslocamento da coluna da água. No primeiro ensaio, o modelo desenvolvido apresentou resultados consistentes na calibração e validação, sendo potencialmente viável para a caracterização e pré-seleção de genótipos visando o melhoramento da macaúba. No segundo ensaio, foram desenvolvidos dois algoritmos, um considerando cada fruto como uma esfera e outro por aproximação a um elipsóide. Os resultados permitem concluir que os algoritmos podem ser usados nas áreas da pré e pós-colheita e na seleção de genótipos em programas do melhoramento da macaúba.

## ABSTRACT

MATSIMBE, Sofrimento Fenias Savanto, M. Sc., Universidade Federal de Viçosa, August of 2012. **Using optical characteristics to estimate the oil content and volume of mesocarp in Macaw palm fruits.** Adviser: Sérgio Yoshimitsu Motoike. Co-Advisers: Francisco Assis de Carvalho Pinto, Hélio Garcia Leite and José Antonio Saraiva Grossi.

The macaúba [*Acrocomia aculeata* (Jacq.) Lodd. Ex Mart.] is a rustic oilseed, with high productivity and multiple potential, which is strongly demanded in the food, cosmetic, and especially the energy industries, due of its suitability for biodiesel production. However, despite the enormous potential, the exploitation of macaw palm still boils down to foraging. Studies aiming at their sustainable use and the resulting domestication of the species are ignorant, because of the lack of practical and efficient methods to support breeding programs on genotypes selections. The main objective of this study was to develop and propose methods to estimate the oil content and volume of the mesocarp in macaw palm fruit using optical characteristics. For methodologies evaluation two assays were carried out. In the first, 420 samples were used to develop calibration model to predict oil content of mesocarp using visible and near infrared spectrometry. The soxhlet method was used as reference. In the second, consisted of 20 samples, were developed imaging algorithms to estimate the mesocarp volume, and as a reference was used water displacement method. In the first assay the developed model showed consistent results in the calibration and validation sets, and it's potentially feasible for preliminary selections and characterization of genotypes intended macaw palm improvement. In the second, were developed two algorithms, one considering each fruit as sphere and another one using ellipsoid approximation. The results show that the algorithms can be used in the pre and post-harvest sector and to make genotypes selection in macaw palm breeding programs.



## 1 – INTRODUÇÃO GERAL

Nos últimos anos, o mundo tem experimentado uma significativa e crescente demanda por óleos vegetais, alavancado principalmente pelo aumento de consumo para fins alimentares em países emergentes, e pelo surgimento de novas aplicações industriais para o óleo vegetal, destacando-se a transformação em biodiesel (IEA Bioenergy, 2009). Segundo previsões da FAO (2011), essa demanda deverá crescer até o ano de 2020 na ordem de 2,2% ao ano, mantendo os preços dessa *commodity* em alta no mercado internacional. Da mesma forma a demanda por óleo vegetal para a produção do biodiesel deverá crescer 60% no mundo, elevando o consumo dos atuais 10% da produção global de óleo para 16% do total produzido no mundo.

No Brasil, o consumo de óleos e gorduras de origem animal e vegetal aumentou consideravelmente desde a implementação do Programa Nacional de Produção e Uso do Biodiesel (PNPB) em 2004. Desde então, os preços da matéria prima tem-se elevado a níveis incompatíveis ao do mercado de combustíveis.

Presentemente, embora o Brasil possua grande diversidade de espécies vegetais oleaginosas, em função do grande desenvolvimento tecnológico e logística de produção, a soja [*Glycine max* (L.) Merr.] é a principal fonte do óleo vegetal, cobrindo 87% da necessidade de matéria prima para a produção do biodiesel (AGRIANUAL, 2012). Todavia, ela apresenta uma baixa produtividade, na ordem dos 0,2 a 0,4 toneladas de óleo/ha (Lopes *et al.*, 2007), insuficiente para sustentar as projeções da demanda para a produção do biodiesel para os próximos anos. Para fazer face a esta crescente demanda faz-se necessário a adoção de espécies mais produtivas. Nesse aspecto, dendê (*Elaeis guineensis* Jacq.) e macaúba [*Acrocomia aculeata* (Jacq.) Lodd. Ex Mart.] ambas com 4,0 toneladas de óleo/ha (AGRIANUAL, 2012) são as espécies com maior potencial para incrementar a sustentabilidade da produção do biodiesel.

O dendê caracteriza-se pela distribuição restrita em função das suas exigências edafoclimáticas, mais de 1800 mm de precipitação anual, acima de 85% de umidade relativa e temperaturas mínimas médias entre 22 – 24°C e as máximas entre 29 – 33°C (Corley e Tinker, 2003), confinando a sua exploração a região da Amazônia e ao longo da costa baiana. Diferentemente, a macaúba pode desenvolver-se em diferentes ambientes, desde o subtropical ao semiárido (Moura *et al.*, 2009;

Arkcoll, 1990), incluindo as regiões de grande sazonalidade hídrica. Esta rusticidade permite uma ampla exploração ao longo do território brasileiro.

Contudo, apesar do enorme potencial, a exploração da macaúba ainda resume-se ao extrativismo. Os estudos visando a sua domesticação e o consequente uso sustentável da espécie são insipientes, dificultando o desenvolvimento de plantios racionais. Em 2007, a Universidade Federal de Viçosa (UFV) deu um grande passo com o desenvolvimento da tecnologia para a produção de sementes pré-germinadas de macaúba, o que possibilitou a produção de mudas em grande escala, até então considerado o gargalo para o desenvolvimento da cultura (Motoike *et al.*, 2007).

Um segundo passo na domesticação é a estruturação de um programa de melhoramento para o desenvolvimento de cultivares de macaúba de alta produção. O melhoramento genético consiste na seleção e na recombinação de genótipos de interesse, que requer acessar as qualidades do fruto, tais como biométricas e químicas.

Os métodos de avaliação biométrica e química adotados atualmente são laboriosos, caros e demorados, o que limita a sua aplicação em grande escala, dificultando os programas de melhoramento genético. Por exemplo, a avaliação de uma única amostra para a determinação do teor de óleo pelo método padrão, o *soxhlet*, requer entre a preparação prévia da amostra e a extração propriamente dita, aproximadamente 50 horas, sem considerar que o solvente (*n*-hexano) utilizado para a extração do óleo é caro, e também, perigoso.

Nesse sentido, a espectrometria do visível e infravermelho próximo (VIS-NIR) para acessar as qualidades químicas e o processamento de imagens digitais para as biométricas constituem ferramenta de avaliação rápida e eficiente, o qual pode ser aplicado em melhoramento genético e na indústria. A espectrometria do visível e infravermelho próximo foi aplicado por Pérez-Vich *et al.* (1998) em girassol, Velasco *et al.* (1999a) em colza, Kim *et al.* (2007) em *Perilla*, Rudolphi *et al.* (2011) e Elfadl *et al.* (2010) em cártamo, Vaknin *et al.* (2011) em pinhão-manso e Quampah *et al.* (2011) em algodão, visando a predição do teor de óleo. Enquanto o processamento de imagem foi aplicado por Bailey *et al.* (2004) e Sabliov *et al.* (2002) em vários produtos agrícolas, Koc (2007) em melão, Rashidi e Seyfi (2008) em *kiwi* e Rashidi *et al.* (2009) em cantaloupe, para acessar as características

biométricas do fruto. Em todos os casos os autores em referência reportaram vantagens comparativas desses métodos, incluindo: economia do tempo, melhor relação custo-benefício, praticidade, repetibilidade, característica não destrutiva e elevada acurácia, além de que são considerados ambientalmente positivos uma vez que não requerem utilização de reagentes químicos.

O objetivo deste trabalho foi desenvolver e propor métodos para estimar o teor de óleo e volume do mesocarpo no fruto da macaúba utilizando características ópticas, através da:

- Calibração do modelo para predição do teor de óleo do mesocarpo usando a espectrometria VIS-NIR;
- Desenvolver algoritmos do processamento de imagens digitais para determinar o volume do mesocarpo.

## **2 – REVISÃO DA LITERATURA**

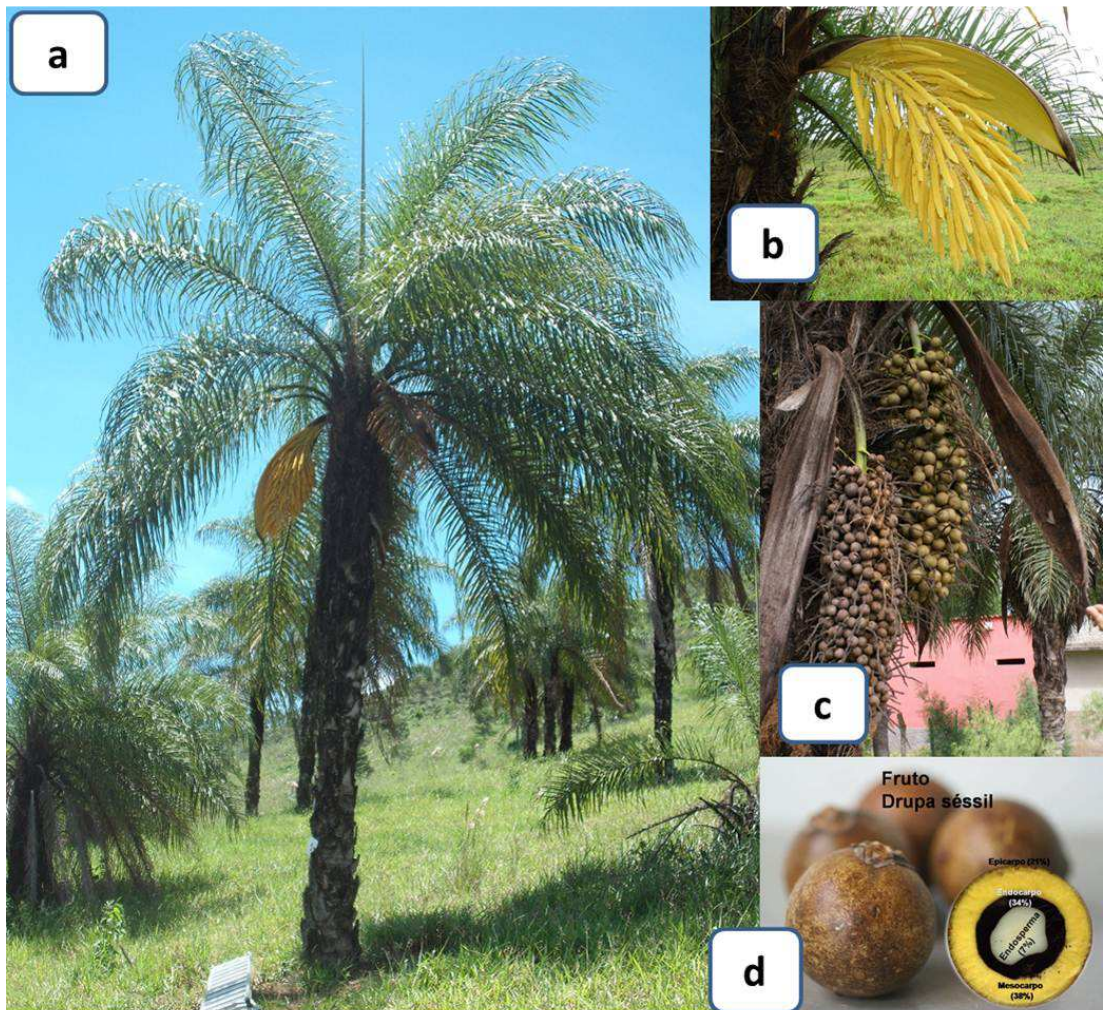
### *2.1 – Aspectos botânicos da macaúba*

A macaúba [*Acrocomia aculeata* (Jacq.) Lodd. Ex Mart.] pertence à família Areaceae, gênero *Acrocomia*. É uma palmeira robusta de crescimento monopodial e de hábito arbóreo (Figura 1a). As folhas frondosas são pinadas e aculeadas. A inflorescência pendente é uma espádice de coloração amarela protegida por espata, que surge nas axilas das folhas (Figura 1b). As flores femininas e masculinas ocorrem em um mesmo cacho, sendo as flores femininas localizadas na base das ráquias e as masculinas na porção apical. O fruto é uma drupa globosa (Figura 1d), constituído por um epicarpo, mesocarpo, endocarpo e de um a três endospermas. O epicarpo é fino e duro, mas rompe-se facilmente quando maduro. O mesocarpo é fibroso, mucilaginoso com coloração amarela ou esbranquiçada. O endocarpo constitui uma camada enegrecida e dura, estando fortemente aderido ao mesocarpo. O endosperma é oleaginoso e revestido de uma fina camada do tegumento (Lorenzi *et al.*, 1996; Henderson *et al.*, 1995; Martins, 1983).

A frutificação da macaúba ocorre, geralmente, na época das chuvas e os frutos atingem a maturação de 12 a 13 meses após antese. Os principais polinizadores são coleópteros das famílias Curculionidae, Nitidulidae e

Escarabaeidae (Scariot, 1998). A inflorescência é também visitada pelas abelhas do gênero *Apis* e abelhas do grupo *Trigonia* (Lorenzi, 2006; Henderson *et al.*, 1995).

A ocorrência da macaúba tem sido apontada para quase toda a América Latina. No Brasil, apresenta uma ampla distribuição com maior concentração nos estados de Minas Gerais, Goiás, Mato Grosso e Mato Grosso do Sul. Consoante à região de ocorrência ela recebe várias denominações: bocaiúva, macacaúba, macaíba, macaibeira, macajuba, macaúba, macaúva, mucajá e mucajaba.



**Figura 1.** Características morfológicas da palmeira macaúba [*Acrocomia aculeata* (Jacq.) Lodd. ex Mart.]: a) vista geral da palmeira; b) inflorescência do tipo espádice; c) cachos com frutos imaturos e maduros; d) vista longitudinal do fruto inteiro e após o corte transversal na qual são destacadas as frações específicas do fruto: epicarpo, mesocarpo, endocarpo e endosperma.

## 2.2 – Utilidades da macaúba

A macaúba é uma planta de alta produtividade e múltiplas potencialidades, sendo o fruto o produto de maior valor econômico. O fruto da macaúba é rico em óleo, estando este concentrado em duas porções: mesocarpo e endosperma.

A concentração de óleo no mesocarpo pode atingir valores superiores a 70%, sendo este de coloração alaranjada e rico em ácido oleico. Ciconini (2012) observou teor do ácido oleico em óleo do mesocarpo de macaúba acima de 80%, o que confere a este o título de “*high oleic*”. Óleos “*high oleic*” são produtos de alto valor e de grande demanda por parte das indústrias alimentícia e, sobretudo, energética, em função da maior estabilidade a oxidação e operabilidade a baixas temperaturas (biodiesel). A concentração do óleo no endosperma é superior a 50%. O óleo extraído do endosperma, ou óleo da amêndoa, é rico em ácidos graxos saturados de cadeia curta, especificamente ácido láurico (Coimbra e Jorge, 2011; Bora *et al.*, 2004; Martins, 1983), constituindo-se em uma valiosa matéria prima para fabricação de sabonetes, shampoos e em outros produtos de saponificação.

As porções do mesocarpo e endosperma também geram tortas e rações, os quais podem ser excelentes suplementos alimentares tanto para humanos e animais, em função da sua rica composição nutritiva. O endosperma, por exemplo, destaca-se como fonte proteica (Hiane *et al.*, 2006a; Martins, 1983) rica em aminoácidos essenciais como valina, metionina+cisteína, isoleucina, e lisina, cujo os valores superam ao padrão teórico da FAO/WHO (1991). Também se destaca pela ausência de proteases resultando numa melhor digestibilidade (Hiane *et al.*, 2006b). No mesocarpo, os nutrientes com maior concentração são os carboidratos (Bora *et al.*, 2004; Martins, 1983),  $\beta$ -caroteno e matéria mineral, essencialmente o potássio, cálcio e fosforo (Ramos *et al.*, 2008).

Outra porção do fruto de grande valor econômico é o endocarpo. O endocarpo da macaúba apresenta elevado poder calorífico e pode ser aproveitado com excelentes rendimentos em gasogênios, operações metalúrgicas e siderúrgicas. Em estudo comparativo dos endocarpos de macaúba e babaçu com a madeira do *Eucalyptus grandis*, Silva (1986) reportou elevados teores da lignina e a alta densidade no endocarpo da macaúba, características que mostram a sua adequação para produção do carvão.

### 2.3 – Métodos analíticos usados para avaliação de produtos agrícolas

Muitos métodos analíticos têm sido propostos para a avaliação de produtos vegetais. Esses métodos são divididos em análises quantitativas ou qualitativas e/ou químicas, sensoriais ou biométricas.

Em oleaginosas, uma das avaliações mais frequente é a quantificação do teor de óleo, na qual o método mais comum e considerado padrão pela AOCS (1994) é o método *soxhlet*. Contudo, apesar de robusto e confiável, este método não se mostra prático, particularmente, para avaliação de amostras numerosas, em função do tempo necessário para o procedimento da análise.

Outros métodos analíticos têm sido propostos para acessar teor de óleo como são os casos da ASE (*Accelerated Solvent Extraction*) e a NMR (*Nuclear Magnetic Resonance*). A ASE emprega a combinação do aumento da temperatura e da pressão, com redução do tempo da preparação da amostra e da quantidade do solvente necessário para completar a extração e, além de manter relativamente constantes as condições da extração. Todavia, é um método trabalhoso e requer uma quantidade substancialmente maior da amostra para operar (Elfadl *et al.*, 2010). A introdução da tecnologia da NMR reduziu significativamente o tempo gasto nas análises, mas não o suficiente para atender as necessidades do mercado atual, visto que, dentre vários aspectos este método requer uma prévia preparação das amostras (Panford e deMan, 1990).

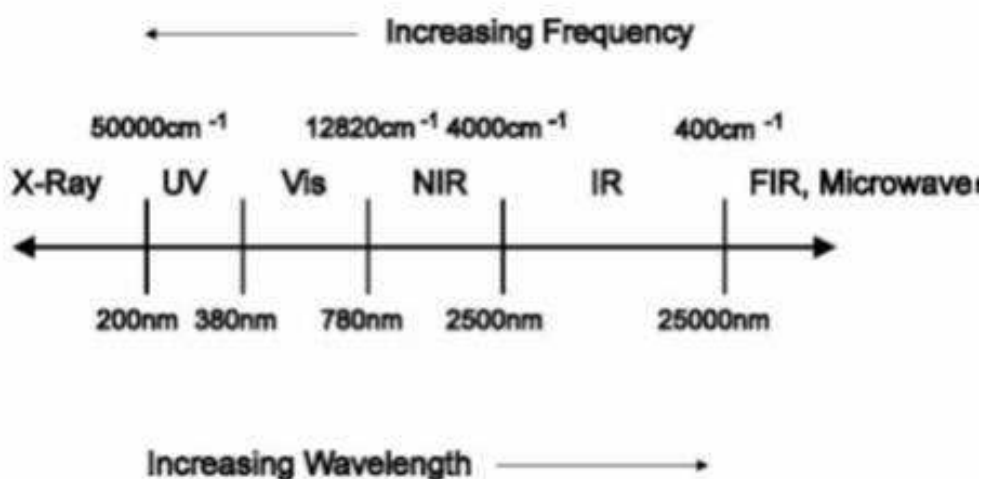
Para a análise das características biométricas, concretamente o volume dos frutos, os métodos convencionais incluem o deslocamento da coluna da água e o deslocamento do gás (Fellegari e Navid, 2011; Rashidi *et al.*, 2009; Rashidi e Seyfi, 2008; Koc, 2007). Todavia, muito embora sejam métodos relativamente simples, em inúmeros casos não são praticáveis, em função da baixa precisão, efeitos adversos aos produtos em análise, e/ou ainda devido a sua limitada operacionalidade (Moreda *et al.*, 2009).

O estágio científico atual exige métodos cada vez mais precisos, elásticos e que sejam econômicos quer em termos de custos ou mesmo do tempo. Este cenário tem propiciado a demanda que se assiste hoje pela utilização da espectrometria do visível e infravermelho próximo (VIS-NIR) e do processamento de imagens digitais como métodos alternativos em diversas avaliações dos produtos agrícolas.

### 2.3.1 – Espectrometria do visível e infravermelho próximo

A espectrometria VIS-NIR faz parte do conjunto dos métodos analíticos que se baseiam na interação da radiação eletromagnética com a matéria, na qual os átomos e moléculas absorvem a energia como função das suas estruturas atômicas (Osborne, 1993). Na região do visível, isto é causado por transições eletrônicas, ou seja, pela transferência dos elétrons para níveis de energia mais elevados mediante a absorção da energia, enquanto na região do infravermelho próximo essa interação entre a matéria e a radiação está relacionada com a absorção de níveis discretos da energia devido a vibrações e rotações das moléculas (Rossel *et al.*, 2005)

A região do visível cobre a faixa espectral com os comprimentos de onda entre 380 a 780 nm (Owen, 2000) e a do infravermelho próximo a faixa entre 780 a 2500 nm (Lin e Ying, 2009; Blanco *et al.*, 1998).



**Figura 2.** Radiação eletromagnética (D.Sathis Kumar *et al.*, 2011)

Atualmente, a espectrometria VIS-NIR constitui alternativa preferencial para melhorar ou substituir os métodos analíticos convencionais nas análises qualitativas e quantitativas em diferentes áreas como agricultura e indústrias alimentar, química, farmacêutica, biomédica e de óleos (Balbin e Smirnov, 2011). Possui diversas vantagens, mas a possibilidade de estimar simultaneamente vários parâmetros numa única análise e de não necessitar da preparação prévia das amostras tem, particularmente, grande relevância quando se trabalha com amostras e análises numerosas (Balbin e Smirnov, 2011; Osborne, 1993). Outra particularidade

importante é a característica não destrutiva que possibilita fazer análises diretamente no campo através da utilização de instrumentos portáteis.

Contudo, quando se trabalha com espectrometria requerem-se processos de calibração bem concebidos de forma a possibilitar o desenvolvimento de melhores modelos para análise dos parâmetros específicos ou de interesse nos produtos em estudo. De acordo com Moghimi *et al.* (2010), os instrumentos de mensuração usados geram um elevado número de dados espectrais que contém ruídos, mas também tem sido difícil distinguir diferenças espectrais sutis nas amostras. Por estas razões é necessária à aplicação da quimiometria para obter conclusões extraíndo informações dos dados espectrais.

A quimiometria usa métodos matemáticos e estatísticos para a seleção dos procedimentos experimentais otimizando a extração da informação química dos sinais espectroscópicos por meio da calibração multivariada (Feudale *et al.*, 2002). Esta técnica inclui a regressão multi-linear (MLR), regressão por componentes principais (PCR) e regressão por mínimos quadrados parciais (PLSR).

MLR permite estabelecer uma correlação entre um reduzido número de comprimentos de onda (variável independente) e uma característica das amostras (variável dependente). Cada comprimento de onda é estudado uma depois da outra e correlacionada com a característica em estudo. Quando a correlação atinge o valor fixado pelo operador é mantida como parte dos comprimentos de onda da calibração do modelo. E finalmente, o modelo é estabelecido entre os comprimentos de onda da calibração e os valores referência da característica em estudo (Roggo *et al.*, 2007).

PCR é uma regressão linear que comporta duas etapas, na primeira os dados espectrais são tratados pela análise por componentes principais (PCA) e na segunda a MLR é realizada na componente principal como variável preditiva (Roggo *et al.*, 2007). PCA envolve procedimentos matemáticos que transformam um número de variáveis correlacionadas em variáveis não correlacionadas denominadas componentes principais.

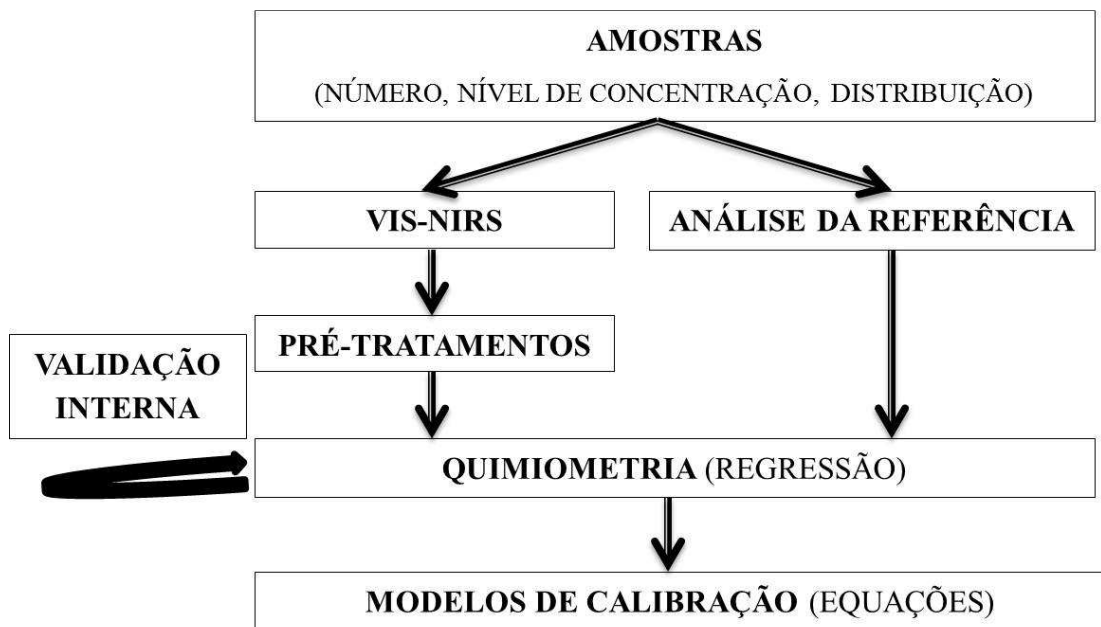
PLSR é um método recente que generaliza e combina as características do PCA e MLR, e que foi desenhado especificamente para lidar com problemas da análise multivariada na qual o número das observações é limitado, há falta de dados e as variáveis estão altamente correlacionadas. Ou seja, quando o objetivo fundamental é a seleção das variáveis relevantes para desenvolver um modelo representativo (Li



*et al.*, 2002). Este método estabelece uma correlação linear entre duas matrizes, os dados espectrais  $X$  e os valores de referência  $Y$ , modelando ambos  $X$  e  $Y$  de forma a encontrar variáveis na matriz  $X$  que melhor descrevem a matriz  $Y$  (Roggo *et al.*, 2007; Yeniay e Göktas, 2002) criando modelos preditivos.

A construção de um modelo envolve: a seleção de amostras para calibração, determinação das propriedades de interesse ( $Y$ ) nas amostras para a calibração, mensuração da resposta espectral ( $X$ ) nas amostras da calibração, calibração do modelo, validação do modelo e aplicação do modelo para a predição (Swierenga *et al.*, 1999).

Em oleaginosas, o método da calibração multivariada mais usado no desenvolvimento de modelos tem sido a PLSR, particularmente na variante MPLSR (mínimos quadrados parciais modificados).



**Figura 3.** Esquema para desenvolvimento de um modelo quantitativo proposto por Roggo *et al.* (2007).

### 2.3.2 – Utilização da espectrometria do visível e infravermelho próximo em oleaginosas

Até à entrada da década de 90 a técnica da espectrometria não era muito aplicada para análise de oleaginosas. Um dos primeiros estudos realizado nesse sentido foi conduzido por Panford *et al.* (1988) no qual verificaram que os

comprimentos de onda de 1230, 1806 e 2170 nm eram os mais adequados para estimar o teor de óleo em sementes do cártamo. Ainda no mesmo período, Panford e deMan (1990) estudaram a determinação do teor de óleo em sementes de 9 diferentes espécies de oleaginosas utilizando a espectrometria NIR. O mesmo estudo, também se estendeu para avaliação da influência do perfil dos ácidos graxos na seleção dos comprimentos de onda relevantes para a predição do óleo.

Do período em referência até ao presente, tem se observado uma demanda crescente de pesquisas objetivando o desenvolvimento de modelos para a predição do teor de óleo usando espectrometria VIS-NIR.

Kohel (1998) avaliando sementes do algodão estabeleceu equações para predição do teor de óleo baseadas na tecnologia NIR ( $R^2 = 0,89 - 0,92$ ). Todavia, estas equações não apresentaram resultados satisfatórios quando avaliou acessos das outras variedades ( $R^2 = 0,00 - 0,84$ ).

Pérez-Vich *et al.* (1998) estudando o potencial da espectrometria NIR para estimar o teor de óleo e o perfil dos ácidos graxos em sementes do girassol, concluíram que é uma técnica confiável e de elevada acurácia ( $R^2$  da validação entre 0,97 a 0,99). O bom desempenho desta tecnologia na estimativa do teor de óleo e perfil dos ácidos graxos foi também verificado por Velasco *et al.* (1999b) em sementes do germoplasma das Brassicaceae,  $R^2 = 0,97$  para o teor de óleo e  $R^2 = 0,93$ ,  $R^2 = 0,95$  e  $R^2 = 0,94$  para as concentrações de  $C_{18:1}$ ,  $C_{18:3}$  e  $C_{22:1}$ , respectivamente. Nos dois trabalhos o método quimiométrico usado foi o MPLSR.

Baye e Becker (2004) estabeleceram um método analítico para avaliar a qualidade ( $R^2 = 0,71$  para teor de óleo e  $R^2 = 0,86$  para proteínas) das sementes intactas da *Vernonia galamensis* combinando a espectrometria NIR e MPLSR. Em sementes do cártamo Rudolphi *et al.* (2005) usaram a tecnologia NIR para desenvolver modelos para estimar o teor de óleo com erro padrão para a validação cruzada igual a 0,90 e  $R^2$  elevado (0,96).

Jiang *et al.* (2007) trabalhando com milho desenvolveram uma metodologia não-destrutiva para determinar os teores da proteína, amido e óleo usando a espectrometria NIR e PLSR. Os coeficientes da validação cruzada e externa foram de 0,91 e 0,94 para o teor da proteína, 0,90 e 0,89 para o teor do amido e 0,94 e 0,95 para o teor de óleo, respectivamente.

Recentemente, também em sementes do cártamo, Elfadl *et al.* (2010) desenvolveram modelos baseados na tecnologia NIR combinada com PLSR para estimar o teor de óleo, tendo estimado o óleo ( $R^2 = 0,90$ ) na região espectral com comprimento de onda entre 1000 e 2175 nm. Em estudo similar, Rudolphi *et al.* (2011) desenvolveram modelos com os quais estimaram o teor de óleo com  $R^2$  elevado (0,91).

### 2.3.3 – Processamento de imagens digitais e sua utilização na estimativa do volume de produtos agrícolas

Processamento de imagens digitais consiste em um conjunto de técnicas para a manipulação de imagens digitais no sentido de facilitar a extração de informações (IBGE, 2001). Este procedimento visa melhorar a informação pictórica da imagem para interpretação humana (Marques Filho e Vieira Neto, 1999), empregando métodos de baixo nível e métodos de alto nível. Os métodos de baixo nível envolvem operações como a redução de ruídos, o aumento do contraste, a extração de bordas e a compressão de imagens. Os métodos de alto nível envolvem tarefas como segmentação da imagem em regiões ou objeto de interesse, descrição desses objetos de forma a reduzi-los a uma forma apropriada para representar o conteúdo da imagem (Pedrini, 2008).

A utilização do processamento de imagens digitais engloba diversas áreas, dentre elas a agricultura, onde é uma importante ferramenta para subsidiar os programas do melhoramento, as áreas da pré e pós-colheita e na área de proteção das plantas. Na indústria de frutos é aplicado especialmente na inspeção da qualidade e determinação do tamanho (Moreda *et al.*, 2009).

Nos últimos anos, tem se observado um interesse crescente pela utilização do processamento de imagens digitais para determinação dos atributos físicos dos frutos, resultante das inúmeras vantagens, como a característica não destrutiva, praticidade, economia do tempo e elevada acurácia (Fellegrari e Navid, 2011; Khojastehnazhand *et al.*, 2009; Rashidi *et al.*, 2009; Rashidi e Seyfi, 2008; Rashidi e Gholami, 2008; Koc, 2007; Riyadi *et al.*, 2007; Hahn e Sanchez, 2000).

Hahn e Sanchez (2000) desenvolveram um algoritmo para determinação do volume da cenoura integrando as áreas de secções da cenoura e imagens ortonormais. Riyadi *et al.* (2007) estabeleceram uma metodologia para estimar o tamanho do

mamão combinando as técnicas do processamento de imagens e redes neurais, utilizando como técnica da aproximação o perfil centroidal da imagem digital.

Koc (2007), em melão, e Rashidi e Gholami (2008), em *kiwi*, estabeleceram algoritmos do processamento de imagens digitais através da técnica do disco e aproximação a um elipsoide perfeito. Enquanto, Rashidi e Seyfi (2008), em *kiwi*, e Rashidi *et al.* (2009), em cantaloupe, utilizaram somente a técnica do disco. Nessa técnica, para cada imagem segmentada bidimensionalmente (com eixo  $X$  e  $Y$ ) do fruto, assume-se que é composta por elementos retangulares, e revolvendo a altura de cada elemento retangular no eixo  $X$  são obtidos discos cilíndricos com diâmetro  $\Delta Y$ . De seguida, determina-se o volume de cada disco cilíndrico ( $V_i$ ) multiplicando a área transversal do disco ( $A_i$ ) pela espessura do disco ( $\Delta X$ ). Finalmente, integrando o  $V_i$  obtém-se o volume total do fruto (Riddle, 1979; Rashidi e Seyfi, 2008).

Khojastehnazhand *et al.* (2009) utilizaram a técnica do processamento de imagens digitais para determinar o volume e a área da laranja. Para o cálculo do volume a aproximação utilizada baseou-se na geometria axi-simétrica. Ainda na laranja, Fellegari e Navid (2011) utilizaram a aproximação a esfera.

### 3 – REFERÊNCIAS BIBLIOGRÁFICAS

AGRIANUAL. **Biocombustível**. 2012. Pag. 25 – 29.

AMERICAN OIL CHEMIST' SOCIETY (AOCS). **Official methods and recommended practices of the American Oil Chemist's Society**. Fourth edition. Washington: American Oil Chemist's Society Press. DC:AOCS. Cd 3d-63, 1994.

ARKCOLL, D. **New crops from Brazil**: In JANICK, J.; SIMON, E. J. (Eds). **Advances in new crops**. Timber Press, Portland, OR. 1990. Pag. 367 – 371.

BAILEY, D. G.; MERCER, K. A.; PLAW, C.; BALL, R.; BARRACLOUGH, H. **High speed weight estimation by image analysis**. In: "Proceedings of the New Zealand National Conference on Non Destructive Testing" (Eds MUKHOPADHYAY, S. C.; BROWNE, R. F.; GUPTA, G. S.), July 27-29, New Zealand, Anais, 2004.

BALBIN, R. M.; SMIRNOV, S. V. **Variable selection in near-infrared spectroscopy: Benchmarking of feature selection methods on biodiesel data.** *Analytica Chimica Acta*, Elsevier Science B.V, N. 692, 2011. Pag 63 – 72.

BAYE, T.; BECKER, H. C. **Analyzing seed weight, fatty acid composition, oil, and protein contents in *Vernonia galamensis* germoplasma by near infrared reflectance spectroscopy.** *Journal of the American Oil Chemist's Society*, Vol. 81, N. 7, 2004. Pag. 641 – 645.

BLANCO, M.; COELHO, J.; ITURRIAGA, H.; MASPOCH, S.; de la PEZUELA, C. **Near-infrared spectroscopy in pharmaceutical industry.** *Critical Review. The Analyst*, Vol. 123, 1998. Pag. 135 – 150.

BORA, P. S.; ROCHA, R. V. M. **Macaíba palm: fatty and amino acids composition of fruits.** *Ciência Tecnologia. & Alimentos*, Vol. 4, N. 3, ALTAGA, 2004. Pag. 158 – 162.

COCININI, G. **Caracterização de frutos e óleo da polpa de macaúba dos biomas do cerrado e pantanal do estado de Mato Grosso do Sul, Brasil.** Dissertação (Mestrado em Biotecnologia), Universidade Católica Dom Bosco, Campo Grande, Mato Grosso do Sul, 2012. 128 pag.

COIMBRA, M. C.; JORGE, N. **Characterization of the pulp and kernel oils from *Syagrus oleracea*, *Syagrus romanzoffiana*, and *Acrocomia aculeata*.** *Journal of Food Science*, Vol. 76, N. 8, 2011. Pag. 1151 – 1161.

CORLEY, R. H. V.; TINKER, P. B. **The oil palm.** Blackwell Science Ltd, a Blackwell Publishing Company, Fourth edition, 2003. 541 pag.

D.SATHIS KUMAR, H. P. R. A.; N.PARTHIBAN, S.; DAVID BANJI, S. S. G. **Near infra-red spectroscopy – An overview.** *International Journal of Chem Tech Research*, Vol. 3, N. 2, 2011. Pag. 825 – 836.

ELFADL, E.; REINBRECHT, C.; CLAUPEIN, W. **Development of near infrared reflectance spectroscopy (NIRS) calibration model for estimation of oil content in a worldwide Safflower germoplasm collection.** *International Journal of Plant Production* 4 (4), 2010. Pag. 259 – 270.

FAO. **Recent trends and medium-term prospects in the global vegetable oil market.** Food and Agricultural Organization of the United Nations. 10<sup>th</sup> International Conference, “Fat-and-oil Industries – 2011: Ukraine, Kiev, 23-25 November. 2011.

FAO/WHO. **Protein quality evaluation.** Rome, Italy: Food and Agricultural Organization of the United Nations, 1991.

FELLEGARI, R.; NAVID, H. **Determining the orange volume using image processing.** IPCBEE, Vol. 9, 2011. Pag. 180 – 184.

FEUDALE, R. N.; WOODY, N. A.; TAN, H.; MYLES, A. J.; BROWN, S. D.; FERRÉ, J. **Transfer of multivariate calibration models: a review.** Chemometrics and Intelligent Laboratory Systems, Elsevier Science B.V. 64, 2002. Pag. 181 – 192.

HAHN, F.; SANCHEZ, S. **Carrot volume evaluation using algorithms.** Journal Agriculture Engineering Research 75, 2000. Pag. 243 – 249.

HENDERSON, A.; GALEANO, G.; BERNAL, R. **Field guide to the palms of the Americas.** Princeton University, 1995. Pag. 166 – 167.

HIANE, P. A.; MACEDO, M. L. R.; SILVA, G. M.; NETO, J. A. B. **Avaliação nutricional da proteína das amêndoas de Bocaíuva [*Acrocomia aculeata* (Jack.) Lodd.] em ratos wistar em crescimento.** Boletim do Centro de Pesquisa & Processamento de Alimentos, Vol. 24, N. 1, 2006a. Pag. 191 – 206.

HIANE, P. A.; BALDASSO, P. A.; MARAGONI, S.; MACEDO, M. L. R. **Chemical and nutritional evaluation of kernels of Bocaíuva [*Acrocomia aculeata* (Jack.) Lodd.].** Ciência Tecnologia & Alimentos 26(3), 2006b. Pag. 683 – 689.

IBGE (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATISTICA). **Processamento digital de imagens.** Manuais técnicos em Geociências. N. 9, 2001. 88 pag.

IEA BIOENERGY. **Bioenergy – a sustainable and reliable energy source: A review of status and prospects.** Main report, 2009. 108 pags.

JIANG, H. Y.; ZHU, Y. J.; WEI, L. M.; DAI, J. R.; SONG, T. M.; YAN, Y. L.; CHEN, S. J. **Analysis of protein, starch and oil content of single intact kernels**

**by near infrared reflectance spectroscopy (NIRS) in Maize (*Zea mays* L.).** Plant Breeding 126, 2007. Pag. 492 – 497.

KHOJASTEHNZHAND, M.; OMID, M.; TABATABAEFFAR, A. **Determination of orange volume and surface area using image processing technique.** International Agrophysics, 23, 2009. Pag. 237 – 242.

KIM, K. S.; PARK, S. H.; CHOUNG, M. G. **Nondestructive of oil content and fatty acid composition in *Perilla* seeds by near infrared spectroscopy.** Journal of Agricultural and Food Chemistry, 55 (5), 2007. Pag. 1679 – 1685.

KOC, A. B. **Determination of Watermelon volume using ellipsoid approximation and image processing.** Journal of Postharvest Biology, 45, 2007. Pag. 366 – 371.

KOHEL, R. J. **Evaluation of near infrared reflectance for oil content of cottonseed.** The Journal of Cotton Science, Vol. 2, Issue 1, 1998. Pag. 23 -26.

LI, B.; MORRIS, J.; MARTIN, E. B. **Model selection for partial least squares regression.** Chemometrics and Intelligent Laboratory Systems, Elsevier Science B.V, N. 64, 2002. Pag. 79 – 89.

LIN, H.; YING, Y. **Theory and application of near infrared spectroscopy in assessment of fruit quality – A review.** Sensing & Instrumentation for Food Quality 3, Springer Science. 2009. Pag. 130 – 141.

LOPES, J. da S.; ANDRADE, T. C. Q. de; SANTANA, G. C. **Biodiesel: oportunidades e desafios.** Comunicação. Bahia Agric. Vol. 8, N. 1, 2007. Pag. 24 – 27.

LORENZI, G. M. A. C. ***Acrocomia aculeata* (Jacq.) Lodd. Ex Mart. – Arecaceae: Bases para o extrativismo sustentável.** Tese (Doutorado em Ciências Agrárias), Universidade Federal do Paraná, 2006. 146 pags.

LORENZI, H.; SOUSA, H. M.; MEDEIROS-COSTA, J. T.; CERQUEIRA, L. S. C. **Palmeiras do Brasil: exóticas e nativas.** Nova Odessa: Editora Plantarum, 1996. 303 pags.

MARQUES FILHO, O.; VIEIRA NETO, H. **Processamento Digital de Imagens**. Brasport, ISBN 8574520098, 1999. 299 pag.

MARTINS, H. **Produção de combustíveis líquidos a partir de óleos vegetais: Estudo das oleaginosas nativas de Minas Gerais**. Relatório Final do Convenio STI – MIC/CETEC, Vol. 1, 1983. 152 pags.

MOGHIMI, A.; AGHKHANI, M. H.; SAZGARNIA, A.; SARMAD, M. **Vis/NIR spectroscopy and chemometrics for the prediction of soluble solids content and acidity (pH) of kiwifruit**. Biosystems Engineering, Elsevier Ltd. 106, 2010. Pag. 295 – 302.

MOREDA, G. P.; ORTIZ-CAÑAVATE, J.; GARCIA-RAMOS, F. J.; RUIZ-ALTISENT, M. **Non-destructive Technologies for fruit and vegetable size determination – A review**. Review article, Journal of Food Engineering, Vol. 92, Issue 2, 2009. Pag. 119 – 136.

MOTOIKE, S. Y.; LOPES, F. A.; OLIVEIRA, M. A. R.; CARVALHO, M.; JUNIOR, A. Q. Sá. **Processo de germinação e produção de sementes pré-germinadas de palmeiras do gênero *Acrocomia***. Registro de patente: BR n. PI0703108-7, 2007.

MOURA, E. F.; MOTOIKE, S. Y.; VENTRELLA, M. C.; JÚNIOR, A. Q. DE SÁ; CARVALHO, M. **Somatic embryogenesis in Macaw Palm (*Acrocomia aculeata*) from zygotic embryos**. Scientiae Horticulturae 119 (2009), 2009. Pag. 447 – 454.

OSBORNE, B. G.; FEARN, T.; HINDLE, P. H. **Practical NIR spectroscopy with applications in food and beverages analysis**. Longman Scientific and Technical, 1993. Pag. 13 – 34.

OWEN, T. **Fundamentals of modern UV-visible spectroscopy – Primer**. Agilent Technologies. 2000. 133 pag.

PANFORD, J. A.; deMAN, J. M. **Determination of oil content of seed by NIR: Influence of fatty acid composition on wavelength selection**. Journal of the American Oil Chemist's Society 67, 1990. Pag. 473 – 482.



PANFORD, J. A.; WILLIAMS, P. C.; deMAN, J. M. **Analysis of oilseeds for protein, oil, fiber and moisture by near infrared reflectance spectroscopy.** Journal of the American Oil Chemist's Society 65:10, 1988. Pag. 1627 – 1634.

PEDRINI, H.; SCHWARTZ, W.R. **Análise de imagens digitais – Princípios, algoritmos e aplicações.** Editora Thomson Learning. 2008. 508 pag.

PÉREZ-VICH, B.; VELASCO, L.; FERNÁNDEZ-MARTÍNEZ, J. M. **Determination of seed oil content and fatty acid composition in Sunflower through the analysis of intact seeds, husked seeds, meal and oil by near infrared reflectance spectroscopy.** Journal of the American Oil Chemist's Society, Vol. 75, N. 5, 1998. Pag. 547 – 555.

QUAMPAH, A.; HUANG, Z. R.; WU, J. G.; LIU, H. Y.; LI, J. R.; ZHU, S. J.; SHI, C. H. **Estimation of oil content and fatty acid composition in Cotton seed kernel powder using near infrared reflectance spectroscopy.** Journal of the American Oil Chemist's Society, 2011. 9 pags.

RAMOS, M. I. L.; RAMOS FILHO, M. M.; HIANE, P. A.; BRAGA NETO, J. A.; SIQUEIRA, E. M. de A. **Qualidade nutricional da polpa da bocaiúva *Acrocomia aculeata* (Jacq.) Lodd.** Ciência Tecnologia & Alimentos 28 (supl.), 2008. Pag. 90 – 94.

RASHIDI, M.; GHOLAMI, M.; ABBASSI, S. **Cantaloupe volume determination through image processing.** Journal of Agriculture Science and Technology, Vol. 11, 2009. Pag. 623 – 631.

RASHIDI, M.; GHOLAMI, M. **Determination of Kiwifruit volume using ellipsoid approximation and image processing methods.** International Journal of Agriculture & Biology, Vol. 10, N. 4, 2008. Pag. 375 – 380.

RASHIDI, M.; SEYFI, K. **Determination of Kiwifruit volume using image processing.** World Applied Science Journal 3 (2), 2008. Pag. 184 – 190.

REDDLE, D. F. **Calculus and analysis geometry.** Wadsworth Publishing Company. 1979.

RIYADI, S.; HUSAIN, H.; HUSSAIN, A.; MUSTAFA, M. M. **Papaya size grading using centroidal profile analysis of digital image.** 6<sup>th</sup> WSEAS International Conference on Circuits, Systems, Electronics, Control & Signal Processing, Cairo, Egypt, Anais , Dec 29 -31, 2007. pag. 347 – 351.

ROGGO, Y.; CHALUS, P.; MAURER, L.; LEMA-MARTINEZ, C.; EDMOND, A.; JENT, N. **A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies.** Journal Pharmaceutical and Biomedical Analysis, Elsevier Science B.V, N. 44, 2007. Pag. 683 – 700.

ROSSEL, R. A. V.; WALVOORT, D. J. J.; McBRATNEY, A. B.; JANIK, L. J.; SKJEMASTAD, J. O. **Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soils proprieties.** Geoderma, Elsevier Science B.V. 131, 2005. Pag. 59 – 75.

RUDOLPHI, S.; BECKER, H. C.; SCHIERHOLT, A.; WITZKE-EHBRECHT, S.V. **Improved estimation of oil linoleic and oleic acid and seed hull fractions in Safflower by NIRS.** Journal of the American Oil Chemist's Society, 2011. 7 pags.

RUDOLPHI, S.; VON WITZKE-EBRECHT, S.; BECKER, H. C. **Estimation of oil content in safflower by near infrared reflectance spectroscopy.** In: ESENDAL E. (Ed.). **A unique crop for oil, spices and health, consequently a better life for you.** Proc. of the 6th Int. Safflower Conf., Anais, 2005. Pag. 344 – 347.

SABLIOV C. M.; BOLDOR, D.; KEENER, K. M.; FARKAS, B.E. **Image processing method to determine surface area and volume of axisymmetric agricultural products.** International Journal of Food Crop 5, 2002. Pag. 641 – 653.

SCARIOT, A. **Seed dispersal and predation of the palm *Acrocomia aculeata*.** Principes 42, 1998. Pag. 5 – 8.

SILVA, J. C. **Endocarpos de Babaçu (*Orbignya* spp. Mart.) e de Macaúba (*Acrocomia sclerocarpa* Mart.) comparados à madeira de *Eucalyptus grandis* para a produção de carvão vegetal.** Dissertação (Mestrado em Engenharia Florestal), Escola Superior de Agricultura “Luiz de Queiroz” da Universidade de São Paulo, 1986. 112 pags.

SWIERENGA, H.; WEIJER, A. P. de; van WIJIK, R. J.; BUYDENS, L. M. C. **Strategy for constructing robust multivariate calibration models.** Chemometrics and Intelligent Laboratory Systems, Elsevier Science B.V, N. 49, 1999. Pag. 1 – 17.

VAKNIN, Y.; GHANIM, M.; SAMRA, S.; DVASH, L.; HENDELSMAN, E.; EISIKOITCH, D.; SAMOCHA, Y. **Predicting *Jatropha curcas* seed-oil content, oil composition and protein content using near infrared spectroscopy – a quick and nondestructive method.** Industrial Crops and Products, Vol. 34, Issue 1, 2011. Pag. 1029 – 1034.

VELASCO, L.; MÖLLER, C.; BECKER, H.C. **Estimation of seed weight, oil content and fatty acid composition in intact single seeds of rapeseed (*Brassica napus* L.) by near infrared reflectance spectroscopy.** Euphytica 106, 1999a. Pag. 79 – 85.

VELASCO, L.; GOFFMAN, F. D.; BECKER, H. C. **Development of calibration equations to predict oil content and fatty acid composition in Brassicaceae germoplasm by near infrared reflectance spectroscopy.** Journal of the American Oil Chemist's Society, Vol. 76, N. 1, 1999b. Pag. 25 – 30.

YENIAY, Ö.; GÖKTAS, A. **A comparison of partial least squares regression with other prediction methods.** Hacettepe Journal of Mathematics and Statistics, Vol. 31, 2002. Pag. 99 – 111.

## ARTIGO 1

---

### **Predição do teor de óleo do mesocarpo da macaúba usando espectrometria do visível e infravermelho próximo**

### **Prediction of Oil content of macaw palm mesocarp using visible and near infrared spectrometry**

#### **RESUMO**

No presente trabalho, foi desenvolvido um modelo para a predição do teor de óleo do mesocarpo nos frutos de macaúba, usando a espectrometria do visível e infravermelho próximo. Os valores de referência foram determinados pelo método *soxhlet*. O modelo foi calibrado usando dados espectrais do mesocarpo dos frutos de macaúba pela regressão por mínimos quadrados parciais, considerando nove variáveis latentes. Os resultados da série da calibração foram consistentes com os da série da validação. O modelo desenvolvido é viável como ferramenta para pré-seleção de genótipos em programas do melhoramento de macaúba.

**Palavras chave:** *Acrocomia aculeata*, quantificação do óleo, calibração e validação VIS-NIR

#### **ABSTRACT**

In this paper, a model was developed to predict the oil content of mesocarp in macaw palm fruit using the visible and near infrared spectrometry. The reference values were determined by soxhlet method. The model was calibrated using spectral data from the mesocarp of macaw palm fruits by partial least squares regression, considering nine latent variables. The results of calibration set were consistent with the validation set. The model is viable as a tool for genotypes preliminary selection in macaúba breeding programs.

**Keywords:** *Acrocomia aculeata*, oil quantification, VIS-NIR validation and calibration

## 1 – INTRODUÇÃO

A macaúba [*Acrocomia aculeata* (Jacq.) Lodd. Ex Mart.] é uma oleaginosa rústica, de alta produtividade e múltiplas potencialidades, justificadas pelo elevado teor de óleo no fruto. O óleo do mesocarpo, que é rico em ácido oleico (COCININI, 2012; COIMBRA e JORGE, 2011), apresenta grande demanda nas indústrias alimentícia e, principalmente, a energética, em função da adequação para produção do biodiesel. O óleo extraído do endosperma é rico em ácido láurico (COIMBRA e JORGE, 2011; BORA e ROCHA, 2004), constituindo-se em uma valiosa matéria prima na indústria de cosméticos. Outra aplicação de importância econômica é a utilização do carvão do endocarpo em operações metalúrgicas e siderúrgicas

Recentemente, estudos visando o melhoramento genético da espécie têm sido estabelecidos nos principais centros de pesquisa no Brasil. Na avaliação do potencial produtivo dos diferentes genótipos são requeridas análises químicas qualitativas e quantitativas no fruto da macaúba. Entretanto, uma das dificuldades enfrentadas por pesquisadores é a capacidade de acessar estas informações no fruto da macaúba em função dos processos laboriosos e caros dos métodos convencionais de avaliação (QUAMPAH *et al.*, 2011). Por exemplo, a avaliação de uma única amostra para a determinação do teor de óleo pelo método padrão, o *soxhlet*, requer entre a preparação prévia da amostra e a extração propriamente dita, aproximadamente 50 horas, sem considerar que o solvente (*n*-hexano) utilizado para a extração do óleo é tóxico e poluente.

Nesse sentido, a espectrometria do visível e infravermelho próximo (VIS-NIR) constitui uma ferramenta de avaliação rápida e eficiente, o qual pode ser aplicado em melhoramento genético e na indústria. Em vários estudos foi proposta a utilização da espectrometria VIS-NIR para desenvolver modelos para predição do teor de óleo em sementes de diversas oleaginosas: girassol (PÉREZ-VICH *et al.*, 1998), colza (VELASCO *et al.*, 1999a), Brassicaceae (VELASCO *et al.*, 1999b) cártamo (RUDOLPHI *et al.*, 2011; ELFADL *et al.*, 2010), pinhão-manso (VAKNIN *et al.*, 2011) e algodão (QUAMPAH *et al.*, 2011). Os referidos estudos reportaram inúmeras vantagens das técnicas espectrométricas, destacando-se: rapidez, elevada precisão, praticidade, custo-benefício, avaliação de mais de um parâmetro numa única análise, característica não destrutiva, não requerem a utilização de reagentes

químicos e nem preparação prévia das amostras (BAYE e BECKER, 2004; ROSSEL *et al.*, 2005). Contudo, até o presente momento não há relatos da sua aplicação em frações específicas do fruto, como o mesocarpo de palmeiras.

O objetivo do estudo foi desenvolver um modelo para a predição do teor de óleo do mesocarpo nos frutos de macaúba usando a espectrometria VIS-NIR.

## **2 – MATERIAL E MÉTODOS**

### *2.1 – Amostragem*

Foram utilizados 840 frutos de macaúba provenientes de várias palmeiras de populações naturais, dos municípios de Acaiaca (20° 23' 33" S, 43° 07' 31" W e 601 m de altitude) e Piranga (20° 41' 46" S, 43° 15' 28" W e 606 m de altitude) no estado de Minas Gerais. Em cada região, os cachos foram marcados na antese para o acompanhamento do desenvolvimento e maturação dos frutos. Dez cachos (um cacho/planta) foram marcados em cada população, totalizando 20 cachos. A colheita dos frutos foi iniciada aos 314 dias após a antese e foi repetida em intervalos que variaram de 15 a 30 dias, até aos 495 dias após a antese. No total foram realizadas nove colheitas, representando cada colheita um estágio de maturação do fruto. A unidade amostral consistiu de dois frutos por cacho/colheita, resultando num total de 420 amostras. Após a colheita, as amostras foram conservadas num *freezer* a -20 °C até o momento da mensuração dos dados espectrais.

### *2.2 – Mensuração dos dados espectrais*

Os dados espectrais foram mensurados através do espectroradiômetro *Field Spec<sup>®</sup> HandHeld 2<sup>TM</sup> Spectroradiometer*, trabalhando no modo da reflectância, na faixa com os comprimentos de onda de 325 a 1075 nm, num intervalo amostral de 1,5 nm e acurácia de  $\pm 0,1$  nm. Uma placa referência branca, com aproximadamente 100 % de reflectância em todo o espectro, foi usada como o padrão de referência, e o tempo de integração foi otimizado para 1,09 segundos, maximizando a mensuração do sinal espectral sem que haja saturação, de forma a minimizar a ocorrência de espectros ruidosos. Antes de cada mensuração retirou-se uma pequena porção do epicarpo para expor o mesocarpo do fruto, no qual foi mensurada a luz refletida pela amostra, à sonda ótica do espectroradiômetro. O fruto foi colocado dentro de um

recipiente de borracha de coloração preta, com uma abertura para o encaixe da sonda do espectroradiômetro. Deste modo, a luz emitida pela sonda incidiu somente sobre o mesocarpo do fruto sem que se dispersasse, assegurando a obtenção de leituras da reflectância somente no mesocarpo e, por forma a minimizar a ocorrência de ruídos relacionados ao espalhamento da luz. O mesocarpo é um tecido mucilaginoso, desta forma, a sonda do espectroradiômetro foi protegida com um filme de PVC transparente, *fasfilme*, durante a mensuração dos dados espectrais (reflectância).

O espectro da reflectância foi determinado através da relação entre a reflectância da amostra e da placa referência branca. Cada espectro da reflectância registrado foi sempre um espectro médio resultante de duas réplicas por amostra e duas leituras por réplica. Após a mensuração, os espectros da reflectância foram processados através do programa computacional *View Spec Pro*<sup>TM</sup> para obtenção dos valores da reflectância utilizadas no ajuste do modelo VIS-NIR, enquanto as amostras foram novamente conservadas no *freezer* a -20 °C até o momento da determinação dos valores de referência do teor de óleo, pelo método *soxhlet*.

A reflectância também foi mensurada no epicarpo do fruto visando o ajuste do modelo VIS-NIR para predição do teor de óleo do mesocarpo no fruto da macaúba, porém os resultados não foram satisfatórios (dados não mostrados).

### 2.3 – Determinação dos valores de referência

As amostras foram descongeladas a temperatura ambiente e depois despulpadas. O mesocarpo resultante foi cortado em pequenos pedaços, que foram secados em estufa ventilada por 48 horas a 65 °C, depois moídos e finalmente distribuídos em cartuchos de papel de filtro,  $\pm 5$  g de amostra/cartucho. A extração do óleo e, posterior, determinação dos valores de referência foram realizadas pelo método *soxhlet*, considerado o método padrão para quantificação do óleo em produtos gordurosos (AOCS, 1994).

### 2.4 – Ajuste do modelo VIS-NIR

Para o ajuste do modelo, somente foi utilizada a faixa espectral de 400 a 999 nm, devido à ocorrência de ruídos significantes nas faixas espectrais de 325 a 399 nm e 1000 a 1075 nm. Antes do ajuste do modelo, os dados espectrais foram pré-tratados usando a técnica da correção multiplicativa do sinal (MSC) (MOGHIMI *et al.*, 2010;

JIANG *et al.*, 2007; BLANCO *et al.*, 2001). MSC é um tratamento baseado na separação de efeitos multiplicativos e aditivos da dispersão da luz em medições NIR, minimizando as variações espectrais que não são resultado da concentração do analito (BLANCO *et al.*, 2001).

Dois terços das amostras foram utilizados na série de calibração e validação cruzada, e o restante na série de validação externa. Para evitar o sobreajuste do modelo, foi determinado o número ótimo de variáveis latentes (VLs) considerando o valor mínimo da raiz quadrada do erro médio quadrático da validação cruzada (RMSEVC) (BALBIN e SMIRNOV, 2011; MEVIK e CEDERKVIST, 2004; YENIAY e GÖKTAS, 2002). A validação cruzada apenas foi utilizada para a definição do número ótimo de variáveis latentes. O modelo VIS-NIR para a predição do teor de óleo no mesocarpo foi ajustado, correlacionando os dados espectrais pré-tratados pela MSC com os valores de referência, usando a regressão por mínimos quadrados parciais (PLSR) (LIU *et al.*, 2009; MEVIK *et al.*, 2004; LI *et al.*, 2002; FEUDALE *et al.*, 2002) através do *software* R (R DEVELOPMENT CORE TEAM, 2012).

## 2.5 – Análise estatística

Para avaliar a habilidade preditiva dos modelos ajustados, foram usados diferentes parâmetros estatísticos:

a) A acurácia da correlação (*bias*) entre o método referencia e o modelo desenvolvido, dado pelo estimador 1.

$$bias = \frac{\sum_{j=1}^n (y_j - \hat{y}_j)}{n} \quad (1)$$

Onde:  $y_j$  = valor mensurado pelo método *soxhlet* para a amostra  $j$ ,  $\hat{y}_j$  = valor predito pelo modelo (VIS-NIR) para a amostra  $j$  e  $n$  = número das amostras.

b) A raiz quadrada do erro médio quadrático (RMSE) entre o método referencia e o modelo desenvolvido, representado por RMSEC para a série de calibração ou estimação e RMSEP para a série de predição ou validação externa. RMSE é dado pelo estimador 2.



$$RMSE = \sqrt{\frac{\sum_{j=1}^n (y_j - \hat{y}_j)^2}{n}} \quad (2)$$

c) O erro relativo percentual (ER) entre o método referencia e o modelo desenvolvido, dado pelo estimador 3.

$$ER(\%) = \frac{(y_j - \hat{y}_j)}{y_j} \times 100 \quad (3)$$

d) O coeficiente de determinação ( $R^2$ ) entre o método referencia e o modelo desenvolvido, que foi estimado para todas as séries: calibração ( $R^2_c$ ), validação cruzada ( $R^2_{vc}$ ) e predição ou validação externa ( $R^2_p$ ), empregando o estimador 4.

$$R^2 = \frac{[\sum_{j=1}^n (\hat{y}_j - \bar{\hat{y}})(y_j - \bar{y})]^2}{\sum_{j=1}^n (\hat{y}_j - \bar{\hat{y}})^2 \sum_{j=1}^n (y_j - \bar{y})^2} \quad (4)$$

Onde:  $\bar{y}$  = valor médio mensurado e  $\bar{\hat{y}}$  = valor médio predito.

### 3 – RESULTADOS E DISCUSSÃO

#### 3.1 – Variabilidade nas amostras de macaúba

Na Tabela 1 é mostrado o sumário da análise da variabilidade das amostras utilizadas para o ajuste do modelo. Conforme os resultados, a variação do teor de óleo verificada entre as amostras, de 0,30 a 55,67 % na série de calibração e 0,44 a 56,03 % na série da validação, sugere que variabilidade entre as amostras foi apropriada para análise pela espectrometria VIS-NIR.

**Tabela 1.** Sumário da análise da variabilidade das amostras de macaúba utilizadas nas séries de calibração e validação

Parâmetro estatístico	Teor de óleo (%)
<b>Série da calibração (estimativa)</b>	
Máximo	55,67
Média	28,82
Mínimo	0,30

Desvio	14,07
<b>Série da validação (predição)</b>	
Máximo	56,03
Média	27,88
Mínimo	0,44
Desvio	14,94

### 3.2 – Efeito do pré-tratamento dos dados espectrais

O efeito do pré-tratamento dos dados espectrais no ajuste do modelo foi avaliado e os resultados são mostrados na Tabela 2. A utilização dos dados espectrais originais requereu 10 VLs para ajustar o modelo VIS-NIR. O pré-tratamento pela MSC permitiu ajustar o modelo VIS-NIR usando menor número de VLs (9), e com melhor habilidade preditiva, a qual é comprovada pelos valores relativamente baixos da RMSEC (5,917) e RMSEP (7,081) e, valores relativamente altos do  $R^2_c$  (0,8223) e  $R^2_p$  (0,7760). Em espectrometria VIS-NIR existem muitas fontes da variabilidade entre as amostras, além da concentração do analito incluem-se, também, os efeitos físicos como a luz e/ou estado da amostra, que contribuem para a utilização de um elevado número de amostras na série de calibração, resultando em modelos complexos com baixa habilidade preditiva (BLANCO *et al.*, 2001). A utilização de pré-tratamentos, como MSC, possibilita a correção das variações espectrais que não são induzidas devido à concentração do analito (MOGHIMI *et al.*, 2010; BLANCO *et al.*, 2001) e, por conseguinte, melhorando a habilidade dos modelos.

Em função dos resultados mostrados na Tabela 2, o ajuste do modelo VIS-NIR foi realizado utilizando os espectros pré-tratados pela MSC.

**Tabela 2.** Sumário do efeito do pré-tratamento sobre os dados espectrais

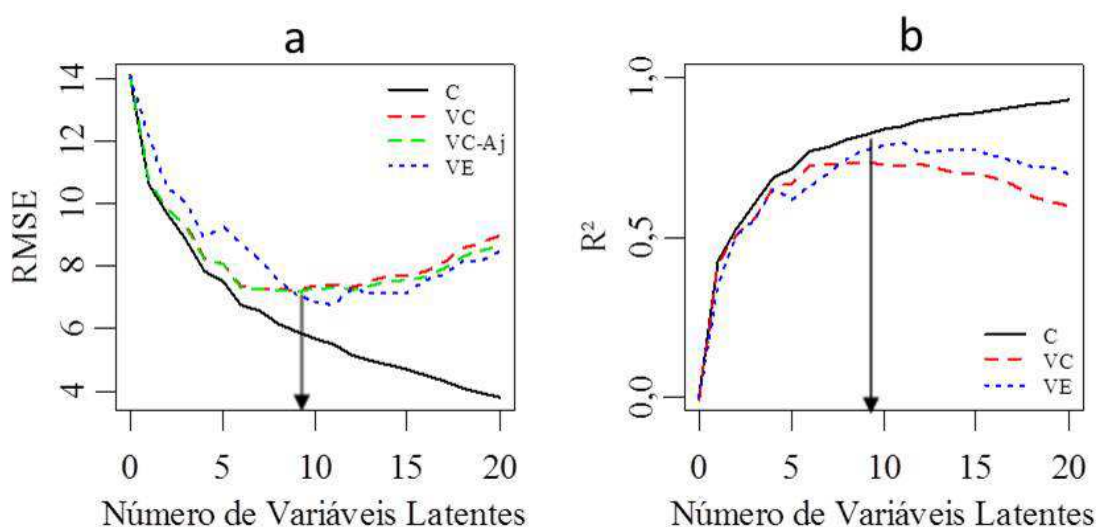
<b>Pré-tratamento</b>	<b>VLs</b>	<b>RMSEC</b>	<b>RMSEP</b>	<b><math>R^2_c</math></b>	<b><math>R^2_p</math></b>
Espectros originais	10	6,592	7,214	0,7796	0,7651
Espectros MSC	9	5,917	7,081	0,8223	0,7760

Onde: *MSC* = correção multiplicativa do sinal, *VLs* = variáveis latentes, *RMSE* = raiz quadrada do erro médio quadrático (*RMSEC* para calibração e *RMSEP* para predição) e  $R^2_c$  = coeficiente de determinação ( $R^2_c$  para calibração e  $R^2_p$  para predição).

### 3.3 – Ajuste do modelo

#### 3.3.1 – Determinação de variáveis latentes

Na Figura 1 é demonstrado que foram necessárias 9 VLs para ajustar o melhor modelo VIS-NIR. Conforme a mesma Figura, o RMSE diminui em todas as séries com o aumento do número de VLs, essa tendência mantém-se até atingir-se 9 VLs. Acima de 9 VLs a tendência do RMSE torna-se divergente nas séries, na calibração e validação externa mantém-se decrescente enquanto na validação cruzada o RMSE aumenta (Figura 1a). O  $R^2$  também mostrou uma tendência divergente nas séries quando se utilizou acima de 9 VLs, tendo sido crescente na série de calibração e decrescente nas séries de validação cruzada e externa. Esse comportamento sugere a ocorrência de modelos sobreajustados, os quais são evitados selecionando o número ótimo de VLs através do RMSEVC mínimo, com o  $R^2$  relativamente alto (QUAMPAH *et al.*, 2011; LIU *et al.*, 2009; SIVAKESAVA e IRUDAYARAJ, 2002; SWIERENGA *et al.*, 1999).



**Figura 1.** Representação gráfica da determinação do número ótimo de VLs utilizado para ajustar o modelo, empregando o RMSE (a) e  $R^2$  (b). Onde: C = calibração, VC = validação cruzada, VC-aj = validação cruzada ajustada, VE = validação externa e VLs = variáveis latentes.

### 3.3.2 - Calibração

Na Tabela 3 estão sumarizadas as estatísticas da série de calibração do modelo VIS-NIR, as quais registraram valores da RMSEC de 5,917,  $bias_c$  de  $-9,2e-15$  e valor do  $R^2_c$  de 0,8223. Esses resultados sugerem uma boa habilidade preditiva do modelo. Quampah *et al.* (2011), Moghimi *et al.* (2010) e Liu *et al.* (2009) consideram que o melhor modelo apresenta valores de  $R^2$  alto e da RMSEC e RMSEP baixos.

**Tabela 3.** Sumário das estatísticas da calibração e validação externa do modelo para predição do teor de óleo do mesocarpo nos frutos de macaúba

Modelo	VL	Modelagem para teor de óleo (%)					
		Calibração (estimação)			Validação externa (predição)		
		$R^2_c$	RMSEC	$bias_c$	$R^2_p$	RMSEP	$bias_p$
VIS-NIR	9	0,8223	5,917	$-9,2e-15$	0,7760	7,081	-0,064

Onde:  $VLs$  = variáveis latentes,  $RMSE$  = raiz quadrada do erro médio quadrático ( $RMSEC$  para calibração e  $RMSEP$  para predição),  $R^2_c$  = coeficiente de determinação ( $R^2_c$  para calibração e  $R^2_p$  para predição) e  $bias$  = acurácia da correlação ( $bias_c$  para calibração e  $bias_p$  para predição).

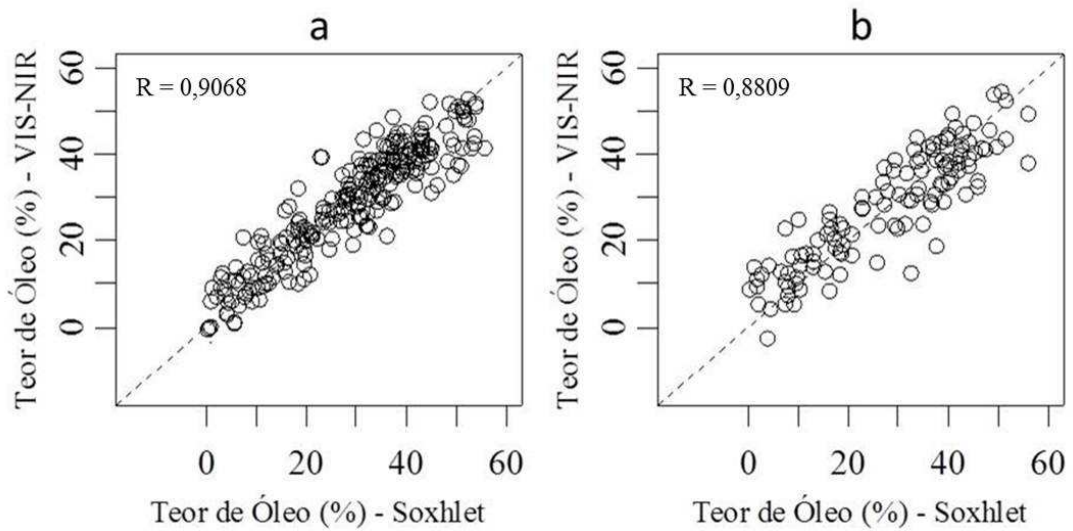
### 3.3.3 – Validação externa

Na mesma Tabela 3, também são apresentados os resultados estatísticos da validação externa, que foram consistentes com os resultados verificados na série da calibração, registrando valores do  $R^2_p$  de 0,7760 e do RMSEP de 7,081. O valor da  $bias_p$  (-0,064) manteve-se próximo do zero, o que indica que a acurácia observada na calibração também foi mantida na série da predição (validação externa). Outro dado relevante verificado foi à elevada consistência (83,56 %) do modelo, mostrando que o mesmo foi otimizado através da inclusão de um número ótimo de VLs. A consistência do modelo é dada pela razão entre o RMSEC e RMSEP.

De acordo com Roggo *et al.* (2003), os modelos quantitativos são considerados com elevada acurácia quando apresentam valores da  $bias$  e RMSE baixos. Enquanto Elfadl *et al.* (2010), considera que para o modelo ser considerado consistente a razão entre RMSEC e RMSEP deve ser próximo do um.

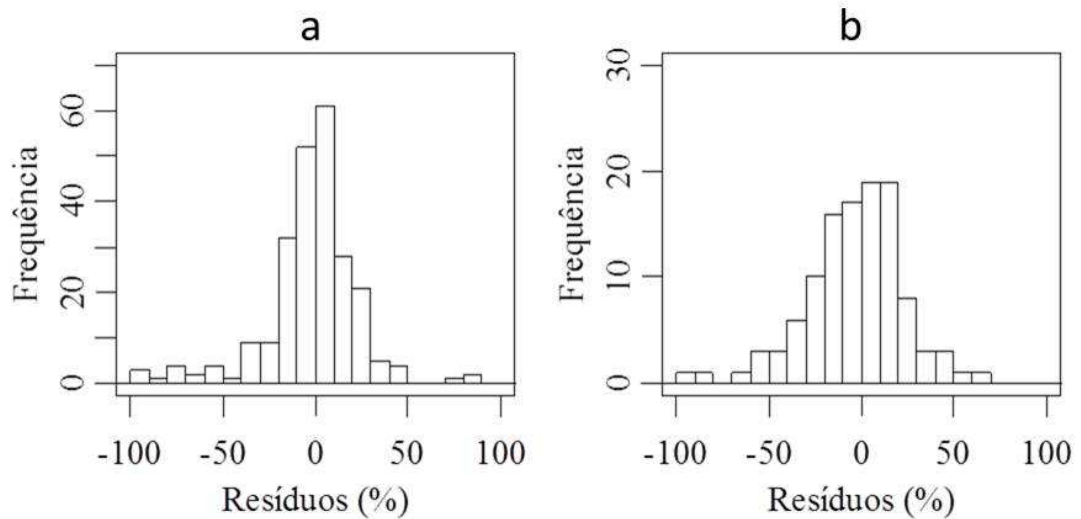
Na Figura 2 é apresentada a visualização gráfica da correlação entre o método *soxhlet* e o modelo VIS-NIR. Nela, pode se observar que existe uma correlação linear entre os valores preditos pelo modelo VIS-NIR com os valores de referência, descrevendo a mesma tendência da linha de 45° na série da calibração (Figura 2a).

A representação gráfica da validação externa também mostra uma correlação significativa entre os valores preditos pelo modelo VIS-NIR com os valores obtidos através do método *Soxhlet* (Figura 2b).



**Figura 2.** Correlação entre os valores estimados e observados pelos dois métodos: calibração (a) e validação externa (b). Onde R = correlação linear.

Os histogramas representados na Figura 3 mostram a distribuição dos resíduos nas séries da calibração e validação externa. Os resíduos estão distribuídos entre valores de  $\pm 100\%$  nas séries, mostrando maior confinamento entre valores de  $\pm 50\%$  e centralização em  $0\%$ , obedecendo a uma distribuição normal.



**Figura 3.** Histogramas dos resíduos: calibração (a) e validação externa (b).

Do modo geral, os resultados verificados foram promissores indicando que a espectrometria VIS-NIR apresenta potencial para ser aplicado como um método alternativo para quantificação de óleo do mesocarpo nos frutos de macaúba.

Um dado interessante prende-se com o fato do modelo VIS-NIR desenvolvido no presente trabalho ter registrado uma robustez de 82 % na calibração e 78 % na predição, na faixa espectral de 400 a 999 nm. Na prática, esse resultado tem grande relevância, considerando que essa faixa coincide com cobertura espectral da maioria dos espectrorradiômetro portáteis. A vantagem prática da utilização dessa faixa está na possibilidade de se fazer as análises tanto no laboratório como, diretamente, no campo.

Os trabalhos conduzidos com diversas oleaginosas objetivando a calibração de modelos para predição do teor de óleo através da espectrometria, aqui referenciados, apenas reportaram resultados na região espectral do infravermelho próximo.

Baye e Becker (2004) estudaram o potencial do NIRS para acessar o perfil dos ácidos graxos e os teores de óleo e da proteína em sementes do germoplasma da *Vernonia galamensis*, tendo verificado o  $R^2$  de 0,71 para o modelo de predição do teor de óleo. A cobertura espectral foi na faixa com comprimento de onda de 1100 a 2500 nm. Elfadl *et al.* (2010) em cártamo ( $R^2 = 0.90$ ) e Quampah *et al.* (2011) no

algodão ( $R^2 = 0,99$ ) foram outros autores que reportaram resultados interessantes trabalhando na faixa espectral com comprimento de onda acima dos 1100 nm.

Em trabalho similar, Velasco *et al.* (1999b), trabalhando com sementes do germoplasma de Brassicaceae, desenvolveram um modelo com  $R^2$  igual a 0,97 para predição do teor de óleo. No entanto, mesmo tendo mensurado a reflectância na faixa espectral entre 400 a 2500 nm, os modelos foram calibrados com informação espectral da faixa entre 1100 a 2500 nm. A informação espectral do comprimento de onda abaixo de 1100 nm foi irrelevante para uma modelagem robusta.

Jiang *et al.* (2007) trabalharam na faixa espectral com comprimento de onda entre 780 a 2500 nm, no entanto apenas as faixas espectral de 1114 a 1836 nm e entre 2171 e 2357 nm foram relevantes para a obtenção de modelos robustos ( $R^2 = 0,94$  a  $0,95$ ) para caracterizar o milho em termos do teor de óleo. Velasco *et al.* (1999a) em colza, embora também tenham trabalhado na faixa espectral entre 400 a 2500 nm, somente as faixas entre 1100 e 1460 nm e entre 1560 e 2500 nm foram relevantes, tendo obtido modelos com  $R^2$  de 0,92. Pérez-Vich *et al.* (1998), a semelhança dos outros autores, verificaram que apenas a faixa espectral entre 1100 a 2500 nm foi relevante para a obtenção de modelos robustos para acessar o teor de óleo nas sementes do girassol. Mas também, reportaram diferenças na robustez dos modelos quando usaram espectros NIRS das sementes intactas ( $R^2$  de 0,76 a 0,85), semente moída ( $R^2$  de 0,92 a 0,98) e semente sem casca ( $R^2$  de 0,90 a 0,97).

Tendo em conta o referenciado, fica evidente que, a par dos atributos químicos, as características físicas das amostras e a região espectral também têm grande relevância na calibração dos modelos preditivos, influenciando a sua robustez.

O modelo VIS-NIR desenvolvido constitui uma alternativa viável, robusta e prática para a quantificação do teor de óleo do mesocarpo no fruto da macaúba. A sua utilização poderá trazer vantagens em termos de economia do tempo e custos, além das vantagens ambientais, uma vez que permitirá reduzir significadamente o número de análises pelo método padrão, o método *soxhlet*.

#### 4 – REFERÊNCIAS BIBLIOGRÁFICAS

AMERICAN OIL CHEMIST' SOCIETY (AOCS). **Official methods and recommended practices of the American Oil Chemist's Society**. Fourth edition. Washington: American Oil Chemist's Society Press. DC:AOCS. Cd 3d-63, 1994.

BALBIN, R. M.; SMIRNOV, S. V. **Variable selection in near-infrared spectroscopy: Benchmarking of feature selection methods on biodiesel data**. *Analytica Chimica Acta*, Elsevier Science B.V, N. 692, 2011. Pag 63 – 72.

BAYE, T.; BECKER, H. C. **Analyzing seed weight, fatty acid composition, oil, and protein contents in *Vernonia galamensis* germoplasma by near infrared reflectance spectroscopy**. *Journal of the American Oil Chemist's Society*, Vol. 81, N. 7, 2004. Pag. 641 – 645.

BLANCO, M.; COELHO, J.; MONTOLIU, I.; ROMERO, M. A. **Orthogonal signal correction in near infrared calibration**. *Analytica Chimica Acta* 434, Elsevier Science B.V, 2001. Pag. 125 – 132.

BORA, P. S.; ROCHA, R. V. M. **Macaíba palm: fatty and amino acids composition of fruits**. *Ciência Tecnologia & Alimentos*, Vol. 4, N. 3, ALTAGA, 2004. Pag. 158 – 162.

COCININI, G. **Caracterização de frutos e óleo da polpa de macaúba dos biomas do cerrado e pantanal do estado de Mato Grosso do Sul, Brasil**. Dissertação (Mestrado em Biotecnologia), Universidade Católica Dom Bosco, Campo Grande, Mato Grosso do Sul, 2012. 128 pag.

COIMBRA, M. C.; JORGE, N. **Characterization of the pulp and kernel oils from *Syagrus oleracea*, *Syagrus romanzoffiana*, and *Acrocomia aculeata***. *Journal of Food Science*, Vol. 76, N. 8, 2011. Pag. 1151 – 1161.

ELFADL, E.; REINBRECHT, C.; CLAUPEIN, W. **Development of near infrared reflectance spectroscopy (NIRS) calibration model for estimation of oil content in a worldwide Safflower germoplasm collection**. *International Journal of Plant Production* 4 (4), 2010. Pag. 259 – 270.



FEUDALE, R. N.; WOODY, N. A.; TAN, H.; MYLES, A. J.; BROWN, S. D.; FERRÉ, J. **Transfer of multivariate calibration models: a review.** Chemometrics and Intelligent Laboratory Systems, Elsevier Science B.V. 64, 2002. Pag. 181 – 192.

JIANG, H. Y.; ZHU, Y. J.; WEI, L. M.; DAI, J. R.; SONG, T. M.; YAN, Y. L.; CHEN, S. J. **Analysis of protein, starch and oil content of single intact kernels by near infrared reflectance spectroscopy (NIRS) in Maize (*Zea mays* L.).** Plant Breeding 126, 2007. Pag. 492 – 497.

LI, B.; MORRIS, J.; MARTIN, E. B. **Model selection for partial least squares regression.** Chemometrics and Intelligent Laboratory Systems, Elsevier Science B.V, N. 64, 2002. Pag. 79 – 89.

LIU, F.; JIANG, Y.; HE, Y. **Variable selection in visible/near infrared spectra for linear and nonlinear calibrations: A case study to determine soluble solids content of beer.** Analytica Chimica Acta 635, Elsevier B.V., 2009. Pag. 45 – 52.

MEVIK, B.-H.; CEDERKVIST, H.R. **Mean Squared Error of Prediction (MSEP) Estimates for Principal Component Regression (PCR) and Partial Least Squares Regression (PLSR).** Journal of Chemometrics 18(9), 2004. Pag. 422 – 429.

MEVIK, B-H; SEGTAN, V. H.; NÆS, T. **Ensemble methods and partial least squares regression.** Journal of Chemometrics 18(11), 2004. Pag. 498 – 507.

MOGHIMI, A.; AGHKHANI, M. H.; SAZGARNIA, A.; SARMA, M. **Vis/NIR spectroscopy and chemometrics for the prediction of soluble solids content and acidity (pH) of kiwifruit.** Biosystems Engineering, Elsevier Ltd. 106, 2010. Pag. 295 – 302.

PÉREZ-VICH, B.; VELASCO, L.; FERNÁNDEZ-MARTÍNEZ, J. M. **Determination of seed oil content and fatty acid composition in Sunflower through the analysis of intact seeds, husked seeds, meal and oil by near infrared reflectance spectroscopy.** Journal of the American Oil Chemist's Society, Vol. 75, N. 5, 1998. Pag. 547 – 555.

QUAMPAH, A.; HUANG, Z. R.; WU, J. G.; LIU, H. Y.; LI, J. R.; ZHU, S. J.; SHI, C. H. **Estimation of oil content and fatty acid composition in Cotton seed kernel**

**powder using near infrared reflectance spectroscopy.** Journal of the American Oil Chemist's Society Journal of the American Oil Chemist's Society, 2011. 9 pags.

R DEVELOPMENT CORE TEAM. **R: A language and environment for statistical computing.** R Foundation for Statistical Computing, ISBN 3-900051-07-0, 2012. [www.R-project.org](http://www.R-project.org).

ROGGO, Y.; DUPONCHEL, L.; RUCKEBUSCH, C.; HUVENNE, J-P. **Statistical tests for comparison of quantitative and qualitative models developed with near infrared spectral data.** Journal of Molecular Structure, Elsevier Science B.V., N. 654, 2003. Pag. 253 – 262.

ROSSEL, R. A. V.; WALVOORT, D. J. J.; McBRATNEY, A. B.; JANIK, L. J.; SKJEMASTAD, J. O. **Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soils proprieties.** Geoderma, Elsevier Science B.V. 131, 2005. Pag. 59 – 75.

RUDOLPHI, S.; BECKER, H. C.; SCHIERHOLT, A.; WITZKE-EHBRECHT, S.V. **Improved estimation of oil linoleic and oleic acid and seed hull fractions in Safflower by NIRS.** Journal of the American Oil Chemist's Society, 2011. 7 pags.

SIVAKESAVA, S.; IRUDAYARAJ, J. **Rapid determination of tetracycline in milk by FT-MIR and FT-NIR spectroscopy.** Journal Dairy Science. 85, 2002. Pag. 487 – 493.

SWIERENGA, H.; WEIJER, A. P. de; van WIJIK, R. J.; BUYDENS, L. M. C. **Strategy for constructing robust multivariate calibration models.** Chemometrics and Intelligent Laboratory Systems, Elsevier Science B.V, N. 49, 1999. Pag. 1 – 17.

VAKNIN, Y.; GHANIM, M.; SAMRA, S.; DVASH, L.; HENDELSMAN, E.; EISIKOITCH, D.; SAMOCHA, Y. **Predicting *Jatropha curcas* seed-oil content, oil composition and protein content using near infrared spectroscopy – a quick and nondestructive method.** Industrial Crops and Products, Vol. 34, Issue 1, 2011. Pag. 1029 – 1034.

VELASCO, L.; MÖLLER, C.; BECKER, H.C. **Estimation of seed weight, oil content and fatty acid composition in intact single seeds of rapeseed (*Brassica***

*napus* L.) by near infrared reflectance spectroscopy. Euphytica 106, 1999a. Pag. 79 – 85.

VELASCO, L.; GOFFMAN, F. D.; BECKER, H. C. **Development of calibration equations to predict oil content and fatty acid composition in Brassicaceae germoplasm by near infrared reflectance spectroscopy.** Journal of the American Oil Chemist's Society, Vol. 76, N. 1, 1999b. Pag. 25 – 30.

YENIAY, Ö.; GÖKTAS, A. **A comparison of partial least squares regression with other prediction methods.** Hacettepe Journal of Mathematics and Statistics, Vol. 31, 2002. Pag. 99 – 111.

## ARTIGO 2

---

### **Estimativa do volume do mesocarpo da macaúba por meio de processamento de imagens digitais**

### **Estimation of volume of macaw palm mesocarp through digital image processing**

#### **RESUMO**

No presente trabalho, foram desenvolvidos algoritmos de processamento de imagens digitais para estimar o volume do mesocarpo nos frutos de macaúba. Os valores de referência foram determinados pelo método do deslocamento da coluna da água. Os algoritmos foram desenvolvidos usando as técnicas do processamento de imagens digitais. O volume do mesocarpo foi estimado considerando cada fruto da macaúba como uma esfera ou um elipsoide. Os algoritmos desenvolvidos são viáveis como ferramenta de avaliação nas áreas da pré e pós-colheita, e para seleção de genótipos em programas do melhoramento da macaúba.

**Palavras chaves:** *Acrocomia aculeata*, Algoritmos, biometria da macaúba, aproximação à esfera, aproximação a elipsóide

#### **ABSTRACT**

In this study, we developed imaging processing algorithms to estimate the mesocarp volume in macaw palm fruit. The reference values were determined by water displacement method. The algorithms were developed using imaging processing techniques. The mesocarp volume was calculated considering each macaw palm fruit as a sphere or an ellipsoid. The developed algorithms are feasible as an assessment tool in the pre and post-harvest sector and for genotypes selection in macaw palm breeding programs.

**Keywords:** *Acrocomia aculeata*, Algorithms, macaw palm Biometrics, the sphere approximation, the ellipsoid approximation

## 1 – INTRODUÇÃO

A macaúba [*Acrocomia aculeata* (Jacq.) Lodd. Ex Mart.] é uma oleaginosa rústica, de alta produtividade e múltiplas potencialidades, justificadas pelo elevado teor de óleo no fruto. O óleo do mesocarpo, que é rico em ácido oleico (COCININI, 2012; COIMBRA e JORGE, 2011), apresenta grande demanda nas indústrias alimentícia e, principalmente, a energética, em função da adequação para produção do biodiesel. O óleo extraído do endosperma é rico em ácido láurico (COIMBRA e JORGE, 2011; BORA e ROCHA, 2004), constituindo-se em uma valiosa matéria prima na indústria de cosméticos. Outra aplicação de importância econômica é a utilização do carvão do endocarpo em operações metalúrgicas e siderúrgicas

Recentemente estudos visando o melhoramento genético da espécie têm sido estabelecidos nos principais centros de pesquisa no Brasil. Nesses estudos, diversos parâmetros biométricos do fruto, tais como a massa, área de superfície e do volume das diferentes partes do fruto são acessados para a determinação do potencial produtivo dos diferentes genótipos. Frequentemente, a avaliação do volume dos frutos tem sido realizada pelos métodos do deslocamento da coluna da água e o deslocamento do gás (RASHIDI *et al.*, 2009; RASHIDI e GHOLAMI, 2008; RASHIDI e SEYFI, 2007). Todavia, muito embora sejam métodos relativamente simples, em inúmeros casos não são praticáveis, em função da baixa precisão (MOREDA *et al.*, 2009), efeitos adversos aos produtos em análise (KOC, 2007), e/ou ainda devido a sua limitada operacionalidade.

O desenvolvimento de recursos computacionais facilitou manipulação de imagens, e conseqüentemente a aplicação das técnicas do processamento de imagens digitais como ferramenta de avaliação eficiente em melhoramento genético, nas áreas da pré e pós-colheita e na indústria de frutos. A aplicação do processamento de imagens como método alternativo para determinar o volume de frutos foi referenciado nos trabalhos de Khojastehnazhand *et al.* (2009) em laranja, Soltani *et al.* (2010) na banana, Koc (2007) no melão, Rashidi e Gholami (2008) em kiwi e Rashidi *et al.* (2009) em cantaloupe. Nesses trabalhos, os autores reportaram vantagens como maior precisão, rapidez e praticidade, em relação aos métodos convencionais. Porém, não há relatos da aplicabilidade das técnicas do

processamento de imagens para determinação do volume de frações do fruto, como o mesocarpo de palmeiras.

O objetivo do estudo foi desenvolver métodos para estimar o volume do mesocarpo de frutos de macaúba por meio do processamento de imagens digitais.

## 2 – MATERIAL E MÉTODOS

### 2.1 – Amostragem

Foram utilizados 20 frutos de macaúba provenientes de várias palmeiras de populações naturais, dos municípios de Acaiaca (20° 23' 33" S, 43° 07' 31" W e 601 m de altitude) e Piranga (20° 41' 46" S, 43° 15' 28" W e 606 m de altitude) no estado de Minas Gerais. Em cada região, os cachos foram marcados na antese para o acompanhamento do desenvolvimento e maturação dos frutos. Cinco cachos (um cacho/planta) foram marcados em cada população, totalizando dez cachos. A colheita dos frutos foi realizada aos 465 e 466 dias após a antese para Acaiaca e Piranga, respectivamente. Em cada cacho foram colhidos dois frutos, um na posição apical e outro na posição basal. A unidade amostral consistiu de um fruto em cada posição do cacho, totalizando 20 amostras que foram utilizados para desenvolver os dois algoritmos do processamento de imagens digitais: modelo de aproximação esférica (MPID-esf) e elíptica (MPID-elip).

### 2.2 – Determinação dos valores de referência

Os valores de referência foram determinados pelo método do deslocamento da coluna da água (MDCA), por imersão de cada fruto numa proveta volumétrica graduada (1000 ml) contendo um volume conhecido de água (volume inicial). Fazendo a diferença do volume final (obtido após a imersão do fruto) pelo inicial obteve-se o volume do fruto ( $V_f$ ). O mesmo procedimento foi repetido após a retirada do epicarpo para obtenção do volume do fruto sem epicarpo ( $V_{se}$ ), e após a retirada do mesocarpo para obtenção do volume do endocarpo contendo o endosperma ( $V_{ee}$ ).

O volume do mesocarpo ( $V_m$ ) foi calculado através da eq. 1.

$$V_m = V_{se} - V_{ee} \quad (1)$$

A proporção média ( $P_m$ ) do volume do mesocarpo em relação ao fruto foi calculada pela eq. 2.

$$P_m (\%) = \sum_{i=1}^n \frac{V_m}{V_f} \times 100 \quad (2)$$

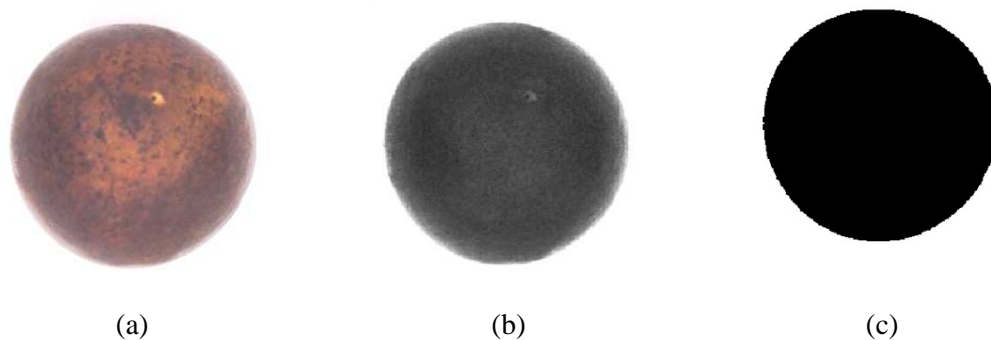
Onde:  $V_m$  = volume do mesocarpo,  $V_{se}$  = volume do fruto sem epicarpo,  $V_{ee}$  = volume do endocarpo contendo endosperma,  $V_f$  = o volume do fruto,  $P_m$  = proporção média do mesocarpo no fruto e  $n$  = número da amostra para cada ensaio, MPID-esf e MPID-elip ( $n = 20$ ).

### 2.3 – Estimativa do volume por meio do processamento de imagens digitais

Foram captadas duas imagens RGB por fruto, uma na posição lateral e outra na posição superior (inserção do pedúnculo) usando uma câmera digital *SONY Cyber-shot*, modelo DSC-W180 de 10.1 *Mega Pixels*, numa câmara de luminosidade uniforme. Para a captação das imagens a câmera digital foi colocada a uma distância de 27 cm paralela ao fruto. As imagens foram processadas e os algoritmos do processamento de imagens digitais desenvolvidos usando o *software* MATLAB 6 RELEASE 12.1 (MATHWORKS, 1996 – 2001).

#### 2.3.1 – Processamento de imagens digitais

O processamento de imagens consistiu na conversão das imagens RGB para a escala de cinza, seguida pela limiarização através do método *Otsu* objetivando a segmentação das regiões de interesse nas imagens (Figura 1). A limiarização consistiu em converter os valores de pixels da imagem, originalmente entre zero e 255, para zero (preto) ou um (branco), com base em um limiar determinado pelo método *Otsu*. Os valores originais de pixels abaixo do limiar foram convertidos para o valor um (cor branca) e os que estiveram acima do limiar foram convertidos para zero (cor preta) (RASHIDI *et al.*, 2009; RASHIDI e GHOLAMI, 2008). Dessa maneira, as imagens segmentadas são binárias (preto e branco). A remoção dos ruídos nas imagens segmentadas foi realizada através de técnicas de morfologia matemática, especificamente as técnicas de preenchimento de regiões (por dilatação) e extração das bordas (por erosão) (PEDRINI e SCHWARTZ, 2008).

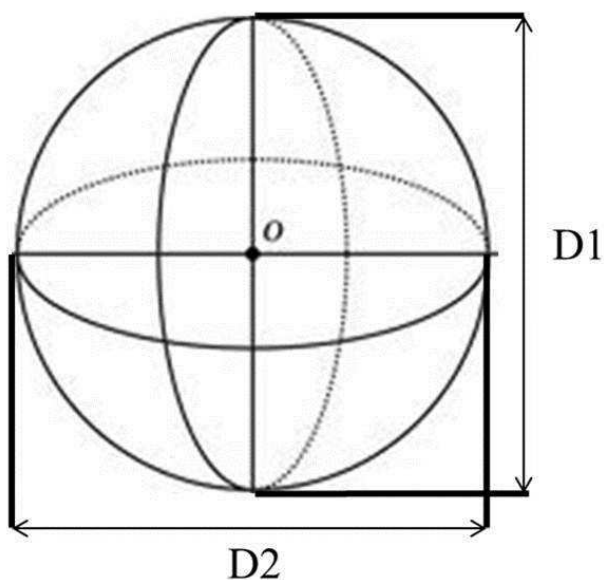


**Figura 1.** Processamento de imagens digitais: fruto da macaúba em RGB (a), escala de cinza (b) e binarizado (c).

### 2.3.2 – Algoritmos do processamento de imagens digitais

#### *Modelo de aproximação esférica (MPID-esf)*

Neste algoritmo, o volume do fruto foi estimado por aproximação do fruto a uma esfera de raio  $r$ , determinado a partir da média dos diâmetros maior  $D1$  e menor  $D2$  (Figura 2). A calibração dimensional foi realizada utilizando uma régua graduada e o *software imageJ* (FERREIRA e RASBAND, 2011) (Figura 3).



**Figura 2.** Dimensões de uma esfera:  $D1$  = diâmetro maior e  $D2$  = diâmetro menor





**Figura 3.** Calibração dimensional utilizando uma régua graduada e o *Software ImageJ* (FERREIRA e RASBAND, 2011) – determinação dos diâmetros em pixel e conversão para as medidas reais do fruto: quantidade de pixel por cm.

O volume estimado do fruto foi calculado usando a eq. 3.

$$V_{ef} = \frac{4}{3} \pi r^3 \quad (3)$$

Onde:  $V_{ef}$  = volume estimado do fruto e  $r$  = raio médio (determinado a partir dos diâmetros maior e menor).

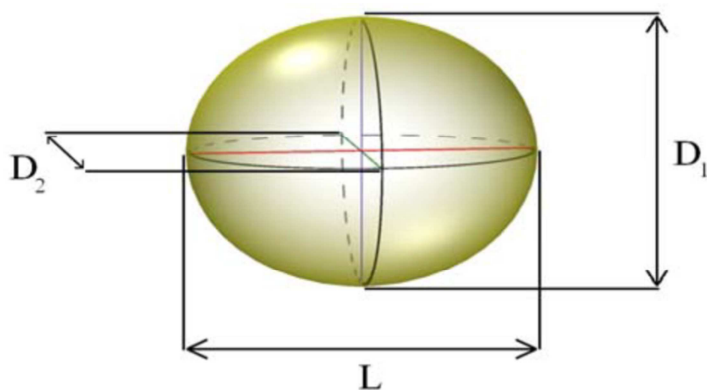
O volume do mesocarpo foi estimado pela eq. 4.

$$V_{em} = V_{ef} \times \frac{P_m}{100} \quad (4)$$

Onde:  $V_{em}$  = é o volume estimado do mesocarpo e  $P_m$  = é a proporção média do mesocarpo no fruto.

#### *Modelo de aproximação elíptica (MPID-elip)*

Este algoritmo foi utilizado para estimar o volume do fruto por a aproximação do fruto a um elipsóide com diâmetro maior  $D_1$ , diâmetro menor  $D_2$  e comprimento  $L$  (Figura 4).



**Figura 4.** Dimensões de um elipsoide (RASHIDI e GHOLAMI, 2008).

O volume estimado do fruto foi calculado através da eq. 6 e o do mesocarpo pela eq. 4.

$$V_{ef} = \pi \frac{L \times D_1 \times D_2}{6} \quad (6)$$

Onde:  $L$  = comprimento,  $D1$  = diâmetro maior e  $D2$  = é o diâmetro menor.

#### 2.4 – Análise estatística

A comparação do volume estimado pelos algoritmos MPID-esf e MPID-elip com o volume determinado pelo MDCA foi realizada através do *Software MedCalc versão 12.3.0.0* (MEDCALC, 1993 – 2012) usando como parâmetros: teste t das amostras dependentes, diferença média entre os métodos, intervalos de confiança para 95%, desvio padrão da diferença entre os métodos, coeficiente de determinação e raiz quadrada do erro médio quadrático. A concordância entre os métodos foi avaliada graficamente através do procedimento proposto por Bland e Altman (1999).

A detecção dos *outliers* foi realizada através dos métodos de Tukey (1977), *Grubbs-double-sided* (GRUBBS, 1969) e teste *Generalized ESD* (ROSNER, 1983), combinados com o teste de *D'agostino Pearson*, importante na verificação da distribuição das amostras. A visualização dos *outliers* foi feita usando os gráficos *Box-and-whisker* (TUKEY, 1977).

A avaliação da identidade dos algoritmos MPID-esf e MPID-elip em relação ao MDCA foi realizada no *Microsoft Excel 2010* usando o procedimento estatístico proposto por Leite e Oliveira (2002). A regra decisória deste teste preconiza que dois métodos são estatisticamente idênticos quando simultaneamente o  $F(H_0)$  e  $t_{\bar{e}}$  não são

significativos e o  $r_{y_j y_1}$  (coeficiente de correlação linear) é maior ou igual  $(1-\bar{e})$  (LEITE e OLIVEIRA, 2002), sendo:

a) Hipótese para avaliação dos algoritmos com MDCA dada pela eq. 7

$$H_0: \beta' = [0 \quad 1] \quad (7)$$

b) Teste F modificado por Graybill (1976) entre os algoritmos e MDCA dado pela eq. 8.

$$F(H_0) = \frac{(\beta - \theta)'(y_1' y_1)(\beta - \theta)}{2RMSE} \sim F_{\infty}(2, n - 2d.f.) \quad (8)$$

c) Erro médio entre os algoritmos e MDCA dado pela eq. 9.

$$\bar{e} = \frac{\sum \frac{y_j - y_1}{y_1}}{n} \quad (9)$$

d) Teste t do erro médio entre os algoritmos e MDCA dado pela eq. 10.

$$t_e = \frac{\bar{e} - 0}{SD_{\bar{e}}} \quad (10)$$

e) Raiz quadrada do erro médio quadrático entre os algoritmos e MDCA dada pela eq. 11

$$RMSE = \sqrt{\frac{\sum_{j=1}^n (y_1 - y_j)^2}{n}} \quad (11)$$

Onde:  $n$  = numero de amostras,  $y_1$  = valor observado através do MDCA,  $y_j$  = valor estimado pelo algoritmo (MPID-esf ou MPID-elip) e  $SD_{\bar{e}}$  = desvio padrão do erro médio.

### 3 – RESULTADOS E DISCUSSÃO

#### 3.1 – Volume observado versus volume estimado

Foram desenvolvidos dois algoritmos MPID-esf e MPID-elip para a estimativa do volume do mesocarpo dos frutos de macaúba. Na Tabela 1 são

mostrados os valores do volume estimados pelos dois algoritmos e os valores obtidos experimentalmente através do MDCA, com os respectivos erros. De uma forma geral, a estimativa do volume pelo algoritmo MPID-esf proporcionou maior erro médio (4,4%) em relação aos valores determinados pelo MDCA, contrastando com o algoritmo MPID-elip, que proporcionou uma estimativa mais próxima dos valores experimentais, demonstrado pelo menor erro médio (1,7 %). A amostra 17 registou o maior erro nas estimativas volume através dos algoritmos, 19,3 % para MPID-esf e 12,3 % para o MPID-elip.

**Tabela 1.** Valores do volume (cm<sup>3</sup>) do mesocarpo dos frutos de macaúba determinados pelo MDCA, MPID-esf e MPID-elip, e os respectivos erros (e).

Amostra	Volumes (cm <sup>3</sup> )		e  (%)	Volumes (cm <sup>3</sup> )		e  (%)
	MDCA	MPID-esf		MPID-elip		
1	32,0	31,3	2,1	31,8	0,6	
2	27,8	27,6	0,7	27,9	0,4	
3	31,2	36,7	17,6	30,8	1,3	
4	34,1	36,5	7,0	33,7	1,2	
5	28,4	29,1	2,4	28,6	0,7	
6	28,6	30,2	5,5	28,9	1,0	
7	31,2	33,5	7,3	31,9	2,2	
8	28,5	28,8	1,0	28,5	0,0	
9	22,1	21,2	4,0	21,8	1,4	
10	25,6	25,8	0,7	25,8	0,8	
11	30,1	30,8	2,3	30,4	1,0	
12	25,3	25,5	0,7	25,5	0,8	
13	26,3	27,0	2,6	27,1	3,0	
14	27,3	27,4	0,3	26,5	2,9	
15	25,0	24,1	3,6	25,1	0,4	
16	28,7	30,1	4,8	28,6	0,3	
17	22,8	27,2	19,3	25,6	12,3	
18	28,2	27,4	2,8	28,7	1,8	
19	26,4	26,5	0,3	26,1	1,1	
20	28,4	28,5	3,5	28,4	0,0	

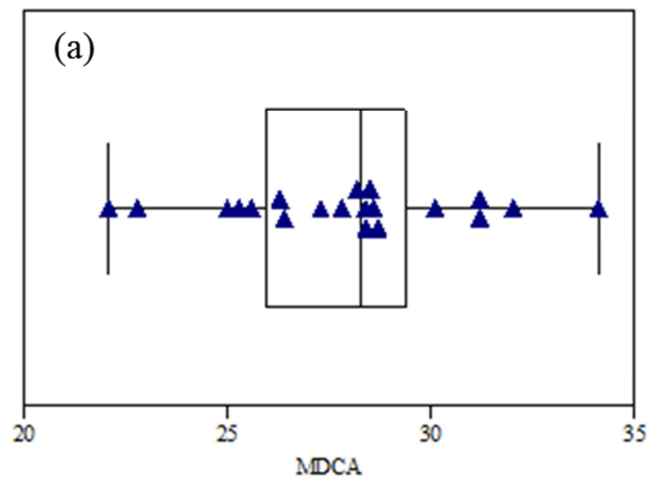
<b>Média</b>	27,9	28,8	4,4	28,1	1,7
--------------	------	------	-----	------	-----

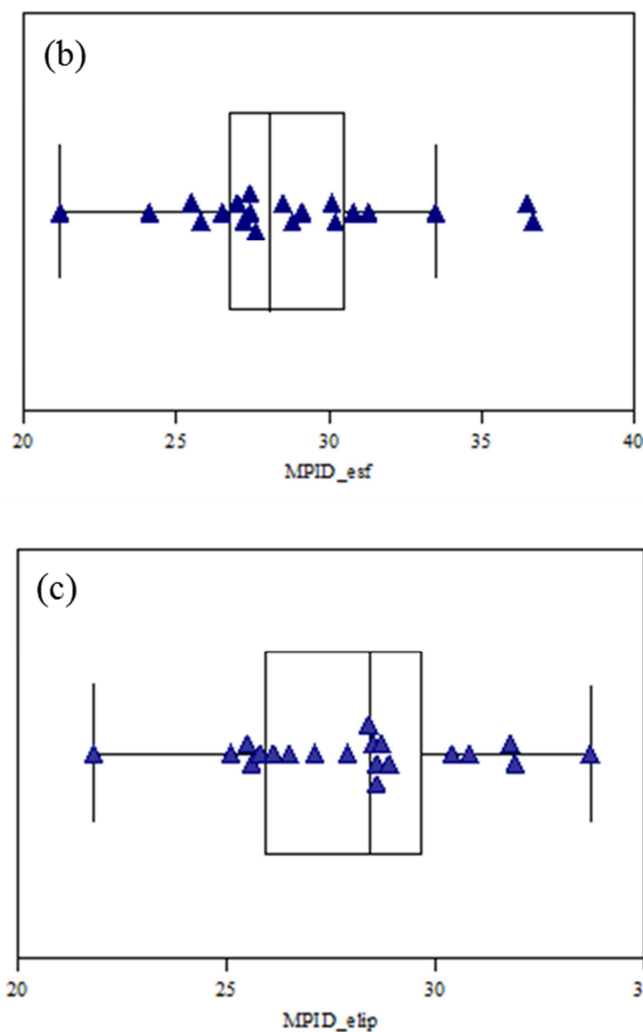
Onde:  $e$  = erro entre os métodos

### 3.2 – Detecção de outliers

A distribuição dos dados do volume e presença de *outliers* nos três métodos foram avaliadas. Pelo teste de *D’Agostino-Pearson*, verificou-se que os dados do volume do MDCA ( $P = 0,9471$ ) e do algoritmo MPID-elip ( $P = 0,8563$ ) apresentam a uma distribuição normal, e não foi detectada a presença de *outliers* pelos métodos de Tukey (1977) e *Grubbs-double-sided* (GRUBBS, 1969), teste *Generalized ESD* (ROSNER, 1983) e da representação gráfica *Box-and-Whisker* (Figuras 5a e 5c).

No algoritmo MPID-esf, através do teste do *D’Agostino-Pearson* também foi verificada uma distribuição normal dos dados do volume ( $P = 0,4012$ ). No entanto, Pelo método de Tukey (1977) detectou-se a presença de dois potenciais *outliers*:  $36,7 \text{ cm}^3$  e  $36,5 \text{ cm}^3$  das amostras 3 e 4, respectivamente. Esses dados foram confirmados como *outliers* através da análise gráfica *Box-and-Whisker* (Figura 5b).





**Figura 5.** Representação gráfica *Box-and-Whisker*, detecção dos *outliers* no conjunto de dados do MCDA (a), MPID-esf (b) e MPID-elip (c).

Ben-Gal (2005) define *outlier* em um conjunto de dados como uma observação que parece ser inconsistente com o restante conjunto. Embora a maioria dos autores considere a rejeição como a forma mais fácil de lidar com os *outliers*, essa opção nem sempre se mostra como a mais válida, pois deve se ter em conta que os *outliers* podem conter informações relevantes para a análise do conjunto em estudo.

A análise aos dois *outliers* do conjunto de dados do volume do algoritmo MPID-esf (método alternativo), integrando o conjunto de dados do MDCA (método padrão) e do algoritmo MPID-elip, mostrou que apenas o dado do volume 36,7 cm<sup>3</sup> da amostra 3 parece ser inconsistente com os demais. Na referida amostra, há um

contraste nos valores do volume dos dois algoritmos em relação ao valor real, o algoritmo MPID-esf regista um erro de 17,6 % em relação ao MDCA, enquanto o algoritmo MPID-elip regista um erro de 1,3 %. Esta é a maior diferença entre os erros dos algoritmos em relação ao método padrão e, um dado discordante atendendo que nas restantes amostras a diferença dos erros dos algoritmos em ao método padrão foi relativamente menor. Desta forma, foi rejeitado o dado do volume da amostra 3 no MPID-esf e, para a comparação desse algoritmo com o MDCA foram consideradas 19 amostras.

### 3.3 – Comparação dos algoritmos com o método padrão

#### 3.3.1 – Comparação do MPID-esf com MDCA

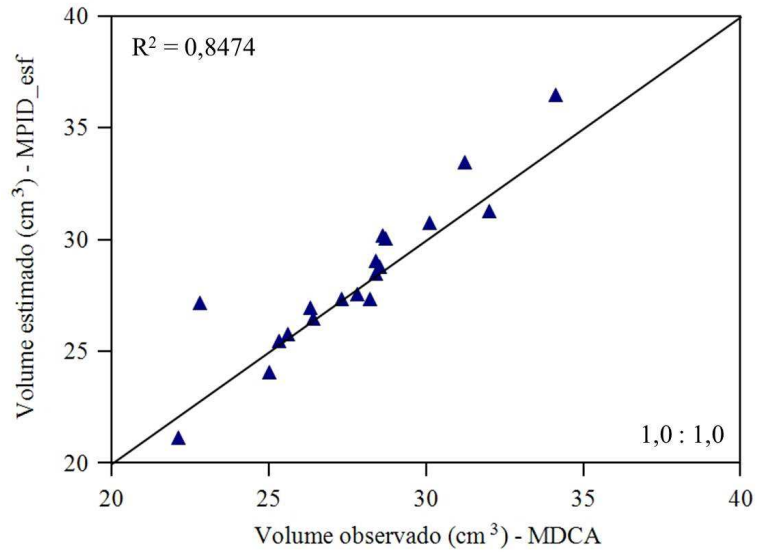
O volume estimado através do algoritmo MPID-esf foi comparado com o volume determinado pelo MDCA, com  $R^2$  de 0,8474 e RMSE de 1,365. Os resultados mostraram que a diferença média ( $d'$ ) verificada entre MDCA e algoritmo MPID-esf foi de  $-0,6 \text{ cm}^3$  e o desvio padrão da diferença de  $1,34 \text{ cm}^3$ . O intervalo de confiança de 95% para avaliação dos métodos foi determinado entre  $-1,26$  e  $0,03 \text{ cm}^3$ . Não foram verificadas diferenças estatísticas entre o volume estimado pelo algoritmo MPID-esf e o volume determinado através do MDCA pelo teste t das amostras dependentes ( $P > 0,05$ ) (Tabela 2).

**Tabela 2.** Sumário dos resultados das estatísticas da comparação dos algoritmos desenvolvidos com o método padrão

Método		n	Volume ( $\text{cm}^3$ )		$R^2$	RMSE	t(H0)
Padrão	Alternativo		$d'$	$SD_{d'}$			P-valor
MDCA	MPID-esf	19	-0,6	1,34	0,8474	1,37	0,060 <sup>ns</sup>
MDCA	MPID-elip	20	-0,2	0,73	0,9418	0,69	0,269 <sup>ns</sup>

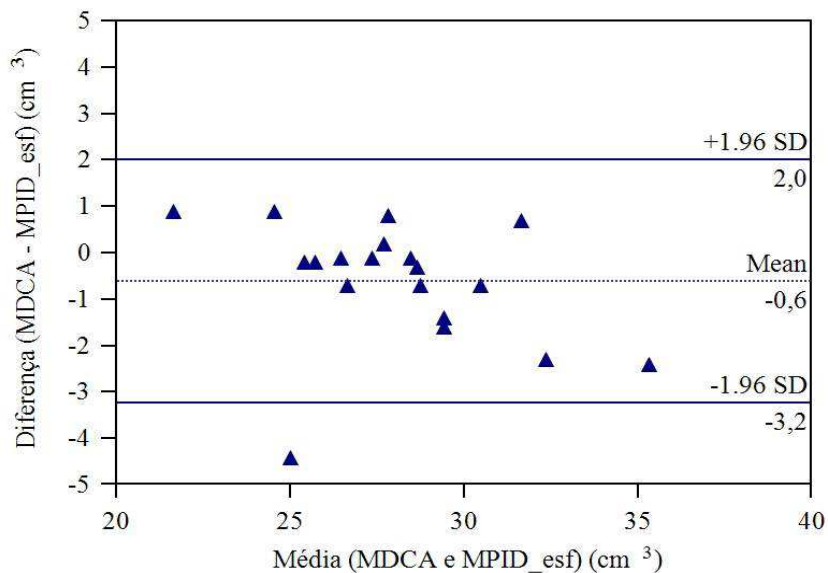
Onde:  $n$  = número de amostras,  $d'$  = diferença média,  $R^2$  = Coeficiente de determinação,  $RMSE$  = Raiz quadrada do erro médio quadrático,  $SD_{d'}$  = Desvio padrão da diferença e  $t(H_0)$  = Teste t das amostras dependentes.

Representação gráfica (1,0:1,0) evidencia a proximidade entre o volume estimado pelo algoritmo MPID-esf e o observado experimentalmente pelo MDCA (Figura 6).



**Figura 6.** Valores do volume observados pelo MDCA *versus* estimados pelo MPID-esf.

Na Figura 7 é demonstrado que as diferenças do volume entre o algoritmo MPID-esf e MDCA foram normalmente distribuídas, sendo esperado que 95% destas diferenças sejam encontradas entre  $-3,2$  e  $2,0$   $\text{cm}^3$ , que são os limites da concordância para a comparação dos dois métodos (BLAND e ALTMAN, 1999). Também, de um modo geral, pode se observar que o algoritmo MPID-esf superestimou o volume do mesocarpo ( $\text{MDCA} - \text{MPID-esf} < 0$ ) (Figura 7).



**Figura 7.** Gráfico de *Bland-Altman* para a comparação do volume do mesocarpo dos frutos da macaúba determinado pelo MDCA e algoritmo MPID-esf, a linha tracejada



na região central do gráfico indica a média das diferenças e as outras duas (regiões de cima e em baixo) indicam os limites de concordância em nível de 95%.

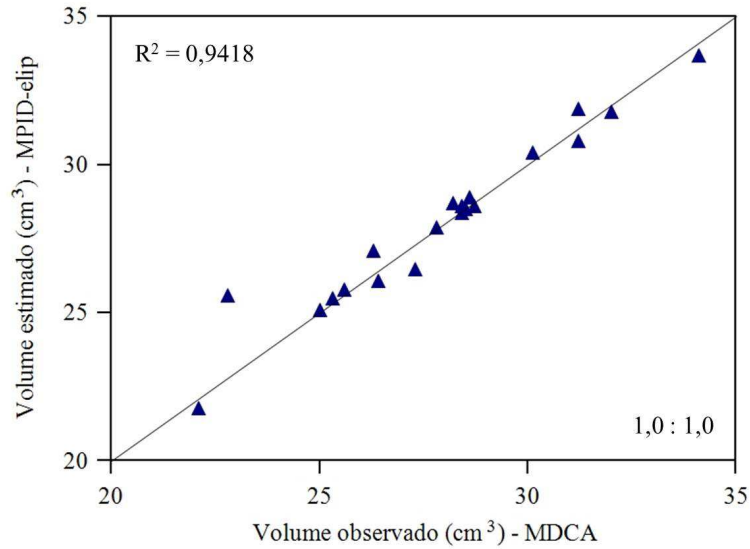
Vários autores relataram efeitos do tamanho do fruto na acurácia dos algoritmos de processamento de imagens digitais desenvolvidos para estimar o volume do fruto. Koc (2007) em melão, Rashidi e Seyfi (2008) e Rashidi e Gholami (2008) em *kiwi* e Rashidi *et al.* (2009) em cantaloupe, verificaram que embora a distância entre a câmera digital e a mesa onde foram captadas as imagens tenha sido constante, a distância entre o fruto e a câmera reduzia com o aumento do tamanho do fruto, resultando numa subestimação do volume em frutos pequenos e superestimação em frutos de tamanhos maiores.

No presente estudo, não foram verificados os efeitos do tamanho do fruto na estimativa do volume (Tabela 1). Porém, deve se levar em conta a uniformidade das amostras utilizadas no estudo em termos do tamanho, o que constitui um dado relevante, pois pode ser que em caso de desuniformidade nas amostras, também no presente estudo, pudessem ser verificados os efeitos do tamanho do fruto relatados em outros estudos.

### 3.3.2 – Comparação do MPID-elip com MDCA

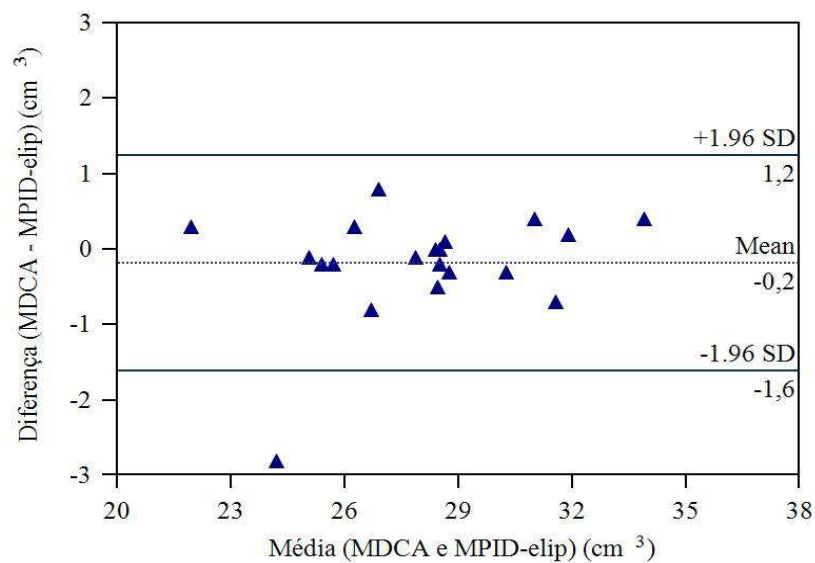
O algoritmo MPID-elip também foi comparado com MDCA e o  $R^2$  de 0,9418 e RMSE de 0,69 foram verificados entre o volume estimado e o observado. Neste caso, os resultados do teste t das amostras dependentes também não mostraram diferenças estatísticas entre o volume do algoritmo MPID-elip e o volume do MDCA ( $P > 0,05$ ). A diferença média ( $d'$ ) entre MDCA e algoritmo MPID-elip foi de  $-0,2 \text{ cm}^3$  e o desvio padrão da diferença foi de  $0,73 \text{ cm}^3$ . O intervalo de confiança de 95% para avaliação dos métodos foi determinado entre  $-0,53$  e  $0,16 \text{ cm}^3$  (Tabela 2).

O diagrama (1,0:1,0) na Figura 8 mostra que houve uma forte correlação entre o volume estimado pelo algoritmo MPID-elip e o observado experimentalmente pelo MDCA.



**Figura 8.** Valores do volume observados pelo MDCA *versus* estimados pelo MPID-elip.

Conforme ilustrado na Figura 9, as diferenças do volume entre o algoritmo MPID-elip e MDCA, também, foram normalmente distribuídas, sendo esperado que 95% destas diferenças sejam encontradas entre  $-1,6$  e  $1,2$   $\text{cm}^3$ , que são os limites da concordância para a comparação dos dois métodos (BLAND e ALTMAN, 1999). Também, neste caso, o algoritmo MPID-elip superestimou o volume do mesocarpo ( $\text{MDCA} - \text{MPID-elip} < 0$ ) (Figura 9) e não foram observados efeitos do tamanho do fruto na acurácia da estimativa do volume do mesocarpo através do algoritmo (Tabela 1).



**Figura 9.** Gráfico de *Bland-Altman* para a comparação do volume do mesocarpo dos frutos da macaúba determinado pelo MDCA e algoritmo MPID-elip, a linha tracejada na região central do gráfico indica a média das diferenças e as outras duas (regiões de cima e em baixo) indicam os limites de concordância em nível de 95%.

Pelos resultados verificados fica evidente que os dois algoritmos desenvolvidos são confiáveis e podem ser usados para estimar o volume do mesocarpo no fruto da macaúba. Porém, importa sublinhar que o algoritmo MPID-elip proporcionou uma estimativa do volume com menor erro em relação ao volume real (observado), e também apresentou maior robustez (94%) dos resultados obtidos. Este melhor desempenho do algoritmo MPID-elip em relação ao MPID-esf pode revelar que, apesar do fruto da macaúba não apresentar uma forma regular, a aproximação a um elipsóide permitiu descrever melhor o formato real do fruto da macaúba.

Não foram encontradas referências de trabalhos envolvendo a aplicação do processamento de imagens digitais para a estimativa do volume de frações do fruto, como o mesocarpo. Ainda assim, os resultados verificados no presente trabalho estão dentro do que tem sido reportado em trabalhos similares com frutos inteiros.

Khojastehnazhand *et al.* (2009) desenvolveram um algoritmo do processamento de imagens digitais para estimar o volume da laranja com  $R^2$  de 0,98, diferença média de  $-0,15 \text{ cm}^3$ , desvio padrão de  $3,41 \text{ cm}^3$  e não observaram diferenças estatísticas entre o algoritmo e o método padrão. Também sem diferenças estatísticas entre os métodos, Soltani *et al.* (2010) em banana, obtiveram  $R^2 = 0,97$ , diferença média de  $1,58 \text{ cm}^3$  e desvio padrão de  $5,51 \text{ cm}^3$ . Koc (2007), em melão, obteve uma diferença média de  $-0,467 \text{ L}$  e desvio padrão de  $0,284 \text{ L}$  usando aproximação a um elipsoide perfeito e diferença média de  $-0,218 \text{ L}$  e desvio padrão de  $0,694 \text{ L}$  usando a técnica de disco. Em cantaloupe, Rashidi *et al.* (2009) verificaram uma diferença média de  $-81,1 \text{ cm}^3$  e desvio padrão de  $237,4 \text{ cm}^3$ . Em kiwi, Rashidi e Gholami (2008) obtiveram uma diferença média de  $-2,23 \text{ cm}^3$  e o desvio padrão de  $8,10 \text{ cm}^3$  usando a técnica do disco. Porém, quando usaram a aproximação a um elipsoide observaram diferenças estatísticas entre os métodos. A diferença média entre os métodos foi de  $6,10 \text{ cm}^3$  e desvio padrão de  $2,97 \text{ cm}^3$ .

No presente trabalho, no algoritmo MPID-elip a diferença média observada foi de  $-0,2 \text{ cm}^3$  e o desvio padrão de  $0,69 \text{ cm}^3$ , enquanto que no MPID-esf obteve-se uma diferença de  $-0,6 \text{ cm}^3$  e desvio padrão de  $1,37 \text{ cm}^3$ , portanto uma diferença média entre os algoritmos e o método padrão bem abaixo da que foi reportada nos trabalhos acima referenciados. Também, o mesmo foi verificado em relação ao desvio padrão.

### 3.4 – Teste de identidade dos algoritmos com o método padrão

Na Tabela 3 está sumarizada a comparação dos algoritmos com o MDCA usando o teste de identidade proposto por Leite e Oliveira (2002).

**Tabela 3.** Sumário dos resultados do teste de identidade dos algoritmos com o método padrão

Método		n	$\bar{e}$	$F(H_0)$	$t_{\bar{e}}$	$r_{Y_j Y_1} \geq (1- \bar{e} )$	Conclusão
Padrão	Alternativo						
MDCA	MPID-esf	19	0,022	2,066 <sup>ns</sup>	1,820 <sup>ns</sup>	não	$Y_j \neq Y_1$
MDCA	MPID-elip	20	0,008	2,181 <sup>ns</sup>	1,145 <sup>ns</sup>	não	$Y_j \neq Y_1$

<sup>ns</sup>  $p > 0,05$

Onde:  $\bar{e}$  = Erro médio,  $F(H_0)$  = Teste  $F$  modificado de Graybill (1976),  $t_{\bar{e}}$  = Teste  $t$  para erro médio,  $r_{Y_j Y_1}$  = Coeficiente de correlação linear,  $Y_j$  = Método alternativo e  $Y_1$  = Método padrão.

Conforme os resultados da Tabela 3, o  $F(H_0)$  e  $t_{\bar{e}}$  entre os algoritmos (MPID-esf e MPID-elip) e o método padrão não foram significativos. O  $r_{Y_j Y_1}$  foi igual a 0,922 e  $(1-|\bar{e}|)$  igual a 0,978 para a comparação do algoritmo MPID-esf com o MDCA, e o  $r_{Y_j Y_1}$  igual a 0,969 e  $(1-|\bar{e}|)$  igual a 0,992 para a comparação do MPID-elip com MDCA. Portanto, em ambos os casos a condição  $r_{Y_j Y_1} \geq (1-|\bar{e}|)$  não foi satisfeita. Em função destes resultados foi rejeitada a identidade dos dois algoritmos com o método padrão.

Nos diversos trabalhos usados como referência a avaliação da exatidão dos algoritmos desenvolvidos com o método padrão foi realizada através do teste t das amostras dependentes, análise da correlação, diferença média e intervalo de confiança e, também, pelo procedimento proposto por Bland e Altman (1999)

(SOLTANI *et al.*, 2010; RASHIDI *et al.*, 2009; RASHIDI e GHOLAMI, 2008; RASHIDI e SEYFI, 2008; KOC, 2007).

Todavia, segundo Leite e Oliveira (2002), a comparação entre métodos analíticos tem sido parcialmente empíricos, já que na maioria dos casos é feita somente analisando a correlação entre dois métodos. Para Abreu *et al.* (1998) (LEITE e OLIVEIRA, 2002), se a correlação entre o método alternativo e o método padrão for elevada, o novo método deve ser adotado. No entanto, é possível verificar uma correlação elevada entre métodos que, ainda assim, gera  $\beta_0$  e  $\beta_1$  muito diferente de zero e um, respectivamente.

De acordo com Leite e Oliveira (2002), nos casos que se comparam dados quantitativos de dois métodos analíticos, o fato comum é a existência de dois vetores de dados quantitativos,  $Y_I$  (método padrão) e  $Y_j$  (método alternativo). Quando se ajusta uma regressão linear  $Y_j = \beta_0 + \beta_1 Y_I + e_i$ , os dois métodos serão estatisticamente iguais se, simultaneamente,  $\beta_0 = 0$  e  $\beta_1 = 1$ . Neste caso os valores de  $R^2$  e  $r_{y_j y_I}$  estarão próximos de um. Desta forma, os mesmos autores propõem um procedimento que permite testar, simultaneamente, se o  $\beta_0 = 0$ , ao mesmo tempo que o  $\beta_1 = 1$  e quantifique o grau de associação entre os métodos. Este Procedimento combina o  $F(H_0)$ ,  $t_{\bar{e}}$  e a relação  $r_{y_j y_I} \geq (1 - |\bar{e}|)$ .

Em termos práticos, o resultado da aplicação do teste de identidade de métodos analíticos proposto Leite e Oliveira (2002) evidencia a subjetividade dos procedimentos estatísticos comumente usados para comparar os métodos alternativos com o método padrão, que também foram usados no presente trabalho para comparar os algoritmos MPID-esf e MPID-elip com MDCA, assim como nos trabalhos usados como referência. Ainda assim, tendo em conta os resultados obtidos, o resultado do teste não invalida a potencial utilização dos algoritmos em áreas como a pré e pós-colheita, bem como áreas afins como os da comercialização e processamento. Na componente do melhoramento genético, os algoritmos são suficientemente robustos ( $R^2 = 0,8474$  e  $R^2 = 0,9418$ ) para a separação de famílias geneticamente distintas, portanto, é viável a sua utilização para seleção de genótipos de interesse.

#### **4 – REFERÊNCIAS BIBLIOGRÁFICAS**

ABREU, C. A.; ABREU, M. F.; ANDRADE, J. C.; van RAIJ, B. **Restrictions in the use of correlation coefficients in comparing methods for the determination of**

**the micronutrients in soils.** Communications on Soil Science and Plant Analysis 29, 1998. Pag. 1961 – 1972.

BEN-GAL, I. **Outlier detection.** In: MAIMON, O.; ROCKACH, L. (Eds). **Data mining and knowledge discovery handbook: A complete guid for practitioners and researchers.** Kluwer Academic Publishers, 2005. 16 pags.

BLAND, J. M.; ALTMAN, D. G. **Measuring agreement in method comparison studies.** Statistical Methods in Medical Research, 8, 1999. Pag. 135 – 160.

BORA, P. S.; ROCHA, R. V. M. **Macaíba palm: fatty and amino acids composition of fruits.** Ciência Tecnologia & Alimentos, Vol. 4, N. 3, ALTAGA, 2004. Pag. 158 – 162.

COCININI, G. **Caracterização de frutos e óleo da polpa de macaúba dos biomas do cerrado e pantanal do estado de Mato Grosso do Sul, Brasil.** Dissertação (Mestrado em Biotecnologia), Universidade Católica Dom Bosco, Campo Grande, Mato Grosso do Sul, 2012. 128 pag.

COIMBRA, M. C.; JORGE, N. **Characterization of the pulp and kernel oils from *Syagrus oleracea*, *Syagrus romanzoffiana*, and *Acrocomia aculeata*.** Journal of Food Science, Vol. 76, N. 8, 2011. Pag. 1151 – 1161.

FERREIRA, T.; RASBAND, W. **ImageJ user guide.** IJ 1.45 m. 2011

GRAYBILL, F. A. **Theory and application of the linear model.** Massachusets: Ouxburg Press, 1976. 704 pag.

GRUBBS, F. E. **Procedures for detecting outlying observations in samples.** Technometrics 11, 1969. Pag. 1 – 21.

KHOJASTEHNZHAND, M.; OMID, M.; TABATABAEEFAR, A. **Determination of orange volume and surface area using image processing technique.** International Agrophysics, 23, 2009. Pag. 237 – 242.

KOC, A. B. **Determination of Watermelon volume using ellipsoid approximation and image processing.** Journal of Postharvest Biol., 45, 2007. Pag. 366 – 371.

LEITE, H .G.; OLIVEIRA, F. H. T. **Statistical procedure to test the identity of analytical methods**. Communications on Soil Science and Plant analysis, 2002. 17 pags.

MATHWORKS. **The language of technical computing installation guide for PC**. RELEASE 12.1 (MATLAB 6 Product Family). 1996 – 2001. [www.mathworks.com](http://www.mathworks.com).

MEDCALC<sup>®</sup>. **Statistics for biomedical**. Trial Version (21 – 22 de Julho de 2012). 1993 – 2012. [www.medcalcsoftware.com](http://www.medcalcsoftware.com).

MOREDA, G. P.; ORTIZ-CAÑAVATE, J.; GARCIA-RAMOS, F. J.; RUIZ-ALTISENT, M. **Non-destructive Technologies for fruit and vegetable size determination – A review**. Review article, Journal of Food Engineering, Vol. 92, Issue 2, 2009. Pag. 119 – 136.

PEDRINI, H.; SCHWARTZ, W.R. **Análise de imagens digitais – Princípios, algoritmos e aplicações**. Editora Thomson Learning. 2008. 508 pag.

RASHIDI, M.; GHOLAMI, M.; ABBASSI, S. **Cantaloupe volume determination through image processing**. Journal Agriculture, Science and Technology, Vol. 11, 2009. Pag. 623 – 631.

RASHIDI, M.; GHOLAMI, M. **Determination of Kiwifruit volume using ellipsoid approximation and image processing methods**. International Journal of Agriculture & Biology, Vol. 10, N. 4, 2008. Pag. 375 – 380.

RASHIDI, M.; SEYFI, K. **Determination of Kiwifruit volume using image processing**. World Applied Science Journal 3 (2), 2008. Pag. 184 – 190.

ROSNER, B. **Percentage points for a generalized ESD many-outlier procedure**. Technometrics 25, 1983. Pag. 165 – 172.

SOLTANI, M.; ALIMARDANI, R.; OMID, M. **A new mathematical modeling of banana fruit and comparison with actual values of dimensional properties**. Modern Applied Science, Vol. 4, N. 8, 2010. Pag. 104 – 113.

TUKEY, J. W. **Exploratory data analysis**. Reading, Mass: Addison Wesley Publishing Company, 1977.

## **CONCLUSÕES GERAIS**

A espectrometria VIS/NIR é uma alternativa viável, robusta e prática para a quantificação do teor de óleo. Ela pode ser usada como uma ferramenta para a pré-seleção de genótipos em programas de melhoramento da macaúba.

O processamento de imagens digitais é uma alternativa viável para estimar o volume do mesocarpo nos frutos de macaúba e os algoritmos desenvolvidos podem ser úteis nas áreas da pré e pós-colheita, bem como áreas afins como comercialização e processamento. Também apresenta potencial como uma ferramenta robusta para a seleção de genótipos de interesse em programas de melhoramento da macaúba.