

# RESSALVA

Atendendo solicitação do(a) autor(a), o texto completo desta dissertação será disponibilizado somente a partir de 28/07/2023.

**UNIVERSIDADE ESTADUAL PAULISTA – UNESP  
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS  
CÂMPUS DE JABOTICABAL**

**CLASSIFICAÇÃO DE MAMOEIROS QUANTO AO TIPO  
SEXUAL E CULTIVARES USANDO A ESPECTROSCOPIA  
NO INFRAVERMELHO PRÓXIMO, ANÁLISES  
MULTIVARIADAS E MACHINE LEARNING**

**Thiago Feliph Silva Fernandes  
Engenheiro Agrônomo**

**2021**

**UNIVERSIDADE ESTADUAL PAULISTA – UNESP  
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS  
CÂMPUS DE JABOTICABAL**

**CLASSIFICAÇÃO DE MAMOEIROS QUANTO AO TIPO  
SEXUAL E CULTIVARES USANDO A ESPECTROSCOPIA  
NO INFRAVERMELHO PRÓXIMO, ANÁLISES  
MULTIVARIADAS E MACHINE LEARNING**

**Thiago Feliph Silva Fernandes**

**Orientador: Prof. Dr. Gustavo Henrique de Almeida Teixeira**

Dissertação apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Mestre em Agronomia (Produção Vegetal).

**2021**

Fernandes, Thiago Feliph Silva

F363c      Classificação de mamoeiros quanto ao tipo sexual e cultivares usando a espectroscopia no infravermelho próximo, análises multivariadas e machine learning / Thiago Feliph Silva Fernandes. --Jaboticabal, 2021 83 p.: il., tabs., fotos

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal

Orientador: Gustavo Henrique de Almeida Teixeira

1. Carica papaya L.. 2. Espectroscopia NIR. 3. Sexagem.

4. Quimiometria. 5. Flores.. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

**CERTIFICADO DE APROVAÇÃO**

**TÍTULO DA DISSERTAÇÃO:** CLASSIFICAÇÃO DE MAMOEIROS QUANTO AO TIPO SEXUAL E CULTIVARES USANDO A ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO, ANÁLISES MULTIVARIADAS E MACHINE LEARNING

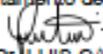
**AUTOR:** THIAGO FELIPH SILVA FERNANDES

**ORIENTADOR:** GUSTAVO HENRIQUE DE ALMEIDA TEIXEIRA

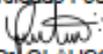
Aprovado como parte das exigências para obtenção do Título de Mestre em AGRONOMIA (PRODUÇÃO VEGETAL), pela Comissão Examinadora:



Prof. Dr. GUSTAVO HENRIQUE DE ALMEIDA TEIXEIRA (Participação Virtual)  
Departamento de Ciências da Produção Agrícola (Produção Vegetal) / FOAV / UNESP - Jaboticabal



Prof. Dr. LUIS CARLOS CUNHA JÚNIOR (Participação Virtual)  
Universidade Federal de Goiás/UFG-Campus Samambaia / Goiânia/GO



Prof. Dr. GLAUCO DE SOUZA ROLIM (Participação Virtual)  
Departamento de Engenharia e Ciências Exatas (DEOEx) / FOAV / UNESP - Jaboticabal

Jaboticabal, 28 de julho de 2021

## DADOS CURRICULARES DO AUTOR

**THIAGO FELIPH SILVA FERNANDES**, nasceu em Castanhal, Pará (PA). Possui graduação em Agronomia, habilitação Engenheiro Agrônomo, pela Universidade Federal Rural da Amazônia (UFRA), Campus Capitão Poço - PA. Durante a graduação foi estagiário voluntário na disciplina de fruticultura no período de 2014 a 2015, realizando pesquisas nas áreas de tecnologia e produção de sementes e mudas de frutíferas nativas e exóticas. Além disso, participou como estagiário voluntário na Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) Amazônia Oriental, no período de 2015 a 2018, auxiliando nas pesquisas de seleção de genótipos de *Citrus* sp. para a região citrícola do Nordeste Paraense. Foi ainda membro grupo de pesquisa Interação Solo-Planta-Atmosfera na Amazônia (ISPAAm) no ano de 2017. Foi bolsista de Iniciação Científica do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) em parceria com EMBRAPA Amazônia Oriental (cota 2018-2019), auxiliando nas avaliações morfológicas e agrônômicas de 12 progênies de *Citrus* sp. para fins de melhoramento genético, bem como na qualidade físico-química dos frutos das mesmas progênies. Em 2019 ingressou no curso de Mestrado acadêmico do Programa de Pós-Graduação em Agronomia (Produção Vegetal), da Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP), Câmpus de Jaboticabal, sob orientação do Prof. Dr. Gustavo Henrique de Almeida Teixeira. Entre 2019 e 2021 desenvolveu pesquisas com o uso da espectroscopia no infravermelho próximo (NIR) visando a classificação de mamoeiros quanto ao tipo sexual e diferentes cultivares. Recentemente foi aprovado no curso de Doutorado do referido Programa de Pós-graduação.

“Eu sou parte de uma equipe.  
Então, quando venço, **não sou eu apenas quem vence.**  
De certa forma termino o trabalho de um **grupo enorme de pessoas.**”

*Ayrton Senna*

A Deus, pelo dom da vida.

À minha família, pelo auxílio e esforço.

Aos meus amigos e professores (principalmente aos da UNESP – Jaboticabal) que tanto contribuíram para minha trajetória acadêmica em especial, ao meu orientador e amigo, Professor Dr. Gustavo Teixeira, peça fundamental na minha formação acadêmica e intelectual.

**Dedico.**

## **AGRADECIMENTOS**

A Deus, Senhor de todas as coisas e destinos, que guia minha trajetória de vida, inspira meu trabalho e não me deixa fraquejar. A ele toda a honra e glória.

À minha mãe Ana Maria N. Silva, por ser minha base e não medir esforços para que eu pudesse realizar meus sonhos, em tantas dificuldades que surgiram no caminho. Aos demais familiares, pois sem o esforço deles, não estaria concluindo mais uma etapa na minha vida.

A todos os professores que cooperaram para minha trajetória acadêmica, contribuindo com a minha formação intelectual e pessoal que, apesar de ser uma profissão importantíssima para a construção de uma sociedade melhor, é tão desvalorizada.

Agradeço ao meu orientador Prof. Dr. Gustavo Henrique de Almeida Teixeira pelos ensinamentos, oportunidades, paciência e preocupação com minha formação

À instituição de ensino Universidade Estadual Paulista (UNESP), Campus Jaboticabal, pela oportunidade de contribuir para minha formação profissional, concedida por meio da realização deste curso.

A todo o corpo docente da instituição UNESP, que repassou um pouco dos seus conhecimentos para agregação da minha formação acadêmica, em especial aos professores que tive durante o mestrado, como Prof. Arthur Bernardes, Prof. Pedro Luiz, Profa. Mara Cristina, Prof. Cristiano Zerbato e Prof. Rouverson Silva que compartilharam seus conhecimentos em disciplinas cursadas durante o mestrado

Aos amigos (a), funcionários do Ripado de Fruticultura, em especial aos funcionários João Brito, Claudemir dos Santos e ao técnico agrícola Adenilson Aparecido Servidone, pois sem eles o desenvolvimento da pesquisa teria encontrado maiores dificuldades.

À Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro - Código Financeiro 148354/2019-0.

Aos amigos: Alex Sanches, Maryelle Barros, Kolima Peña, Jonatan, Antonio Maricélio, Elbys Bastos, Lucas Soares e Valquíria Soares pelos momentos compartilhados com muita alegria.

A todas as pessoas que direta ou indiretamente contribuíram para a realização desta etapa na minha vida acadêmica e pessoal.

**Meus sinceros agradecimentos,  
Thiago Feliph S. Fernandes.**

## SUMÁRIO

<b>RESUMO</b> .....	i
<b>ABSTRACT</b> .....	i
<b>LISTA DE TABELAS</b> .....	i
<b>LISTA DE FIGURAS</b> .....	i
<b>CAPÍTULO 1 - Considerações gerais</b> .....	1
1. Introdução .....	1
2. Revisão de literatura.....	3
2.1 O Mamoeiro .....	3
2.2 Tipos de flores.....	5
2.3 Reprodução e tipos sexuais.....	6
2.4 Herança do sexo .....	7
2.5 Métodos de sexagem.....	8
2.6 Espectroscopia no infravermelho próximo (NIR).....	9
3. Referências .....	17
<b>CAPÍTULO 2 - Determinação sexual em sementes e folhas de mamoeiros utilizando espectroscopia no infravermelho próximo combinadas a técnicas multivariadas e machine learning</b> .....	25
1. Introdução .....	26
2. Material e Métodos.....	27
2.1 Material Vegetal .....	27
2.2 Aquisição dos espectros NIR .....	28
2.3 Método de referência: identificação do tipo sexual .....	29
2.4 Quimiometria.....	29
2.5 Modelos supervisionados.....	31
3. Resultados .....	33
3.1 Sementes.....	33
3.2 Folhas .....	34

4.	Discussão .....	35
4.1	Sementes .....	35
4.1	Folhas .....	36
5.	Conclusão .....	38
6.	Referências .....	38
<b>CAPÍTULO 3 - Classificação de cultivares de mamoeiros por espectroscopia no infravermelho próximo combinada à PLS-DA e machine learning .....</b>		<b>50</b>
1.	Introdução .....	51
2.	Material e Métodos .....	53
2.1	Material Vegetal .....	53
2.2	Aquisição dos espectros NIR .....	53
2.3	Quimiometria .....	54
2.4	Modelos supervisionados .....	55
3.	Resultados .....	57
3.1	Sementes .....	57
3.2	Folhas .....	58
4.	Discussão .....	59
4.1	Sementes .....	59
4.2	Folhas .....	60
5.	Conclusão .....	61
6.	Referências .....	61
<b>CAPÍTULO 4 – Considerações Finais .....</b>		<b>69</b>

# CLASSIFICAÇÃO DE MAMOEIROS QUANTO AO TIPO SEXUAL E CULTIVARES USANDO A ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO, ANÁLISES MULTIVARIADAS E MARCHINE LEARNING

## RESUMO

A sexagem do mamoeiro (*Carica papaya* L.) é uma das práticas hortícolas mais importantes para o sucesso desta cultura, pois esta espécie apresentar flores femininas, hermafroditas e masculinas. Todavia, apenas as plantas hermafroditas produzem frutos com o formato exigido pelo mercado, ou seja, frutos mais alongados e com polpa carnosa. Desta forma, é necessário o plantio de pelo menos três mudas por cova para se garantir a presença de plantas hermafroditas, o que acarreta um maior custo para o produtor. Assim, o objetivo geral deste projeto foi verificar a possibilidade de se utilizar a espectroscopia no infravermelho próximo (NIR) como um método não destrutivo para a sexagem e identificação de cultivares de mamoeiros. Foram utilizados as cultivares 'T2', 'Formosa' e 'Calimosa', do grupo Formosa, e 'THB' e 'Ouro', do grupo Solo. Os espectros NIR foram coletados nas sementes e nas folhas dos respectivos *seedlings*. Os resultados para a classificação do tipo sexual dos mamoeiros utilizando as sementes resultou em um valor de F-score de 0,81 para o grupo de validação externa ao se utilizar a análise de componentes principais e análise discriminante quadrática (PCA-QDA). Para as folhas dos *seedlings* o F-score foi 0,79 usando a PCA e análise discriminante linear (PCA-LDA). Em relação aos resultados para a classificação das cultivares, ao se utilizar as sementes foi possível obter valores de F-score apenas para as cultivares 'THB' (0,85) e 'Ouro' (0,77) no grupo de validação externa com o emprego da regressão por mínimos quadrados parciais e análise discriminante (PLS-DA). Para as demais cultivares não foi possível obter modelos de classificação robustos utilizando tanto as sementes quanto as folhas. Conclui-se que é possível utilizar a espectroscopia NIR associada às técnicas multivariadas como método não destrutivo para a sexagem e classificação de cultivares de mamoeiro utilizando tanto as sementes quanto as folhas dos *seedlings*.

**Palavras-chave:** *Carica papaya* L.; espectroscopia NIR; sexagem; quimiometria; flores.

# CLASSIFICATION OF PAPAYA PLANTS BY SEX AND CULTIVAR USING NEAR INFRARED SPECTROSCOPY, MULTIVARIATE ANALYSIS AND MACHINE LEARNING

## ABSTRACT

Papaya plants can have typical pistillate or female flowers, hermaphrodite flowers and typical staminate or male flowers, but hermaphrodite plants produce fruit with the shape required by the market, that is, more elongated fruit with fleshy pulp. Therefore, it is necessary to plant at least three seedlings per hole to ensure the presence of hermaphrodite plants, which entails a higher cost for the producer. Thus, the general objective of this study was to verify the possibility of using near infrared spectroscopy (NIR) as a non-destructive method for sexing and identification of papaya cultivars. The cultivars 'T2', 'Formosa' and 'Calimosa', from the Formosa group, and 'THB' and 'Ouro', from the Solo group, were used. NIR spectra were collected from the seeds and leaves of the respective seedlings. The results for the classification of the sexual type of papaya trees using the seeds resulted in a F-score value of 0.81 for the external validation group applying principal component analysis and quadratic discriminant analysis (PCA-QDA). For the leaves of the seedling, this value 0.79 using PCA and linear discriminant analysis (PCA-LDA). Regarding the results for the cultivar classification, when seeds were used it was possible to obtain F-score values just for the cultivars 'THB' (0.85) and 'Ouro' (0.77) in the external validation group applying partial least squares regression and discriminant analysis (PLS-DA). For the other cultivars it was not possible to obtain good classification models using both seeds and leaves. It is possible to use NIR spectroscopy associated with multivariate techniques as a non-destructive method for sexing and classification of papaya cultivars using both seeds and leaves of seedlings. However, more studies are needed to improve the performance of the developed models.

**Keywords:** *Carica papaya* L.; NIR spectroscopy, sexing; fraud; flowers.

## LISTA DE TABELAS

### **CAPÍTULO 2 - Determinação sexual em sementes e folhas de mamoeiros utilizando espectroscopia no infravermelho próximo combinadas a técnicas multivariadas e machine learning**

**Tabela 1.** Número de sementes e folhas utilizadas no conjunto de calibração (treinamento), validação cruzada (teste) das cultivares de mamoeiro ‘Calimosa’, ‘Formosa’, ‘T2’, ‘Ouro’ e ‘THB’.....42

**Tabela 2.** Desempenho do conjunto de treinamento, teste e validação externa para o tipo sexual de sementes de mamoeiros. ....43

**Tabela 3.** Matriz de confusão para os conjuntos de validação cruzada (teste) e validação externa do desempenho do modelo PCA-QDA das sementes de todas as cultivares combinadas, separadas de acordo com o tipo sexual do mamoeiro, femininas e hermafroditas. ....44

**Tabela 4.** Desempenho do conjunto de treinamento, teste e validação externa para o tipo sexual de folhas de seedlings de mamoeiros. ....45

**Tabela 5.** Matriz de confusão para os conjuntos de validação (teste) e validação externa do desempenho do modelo PCA-QDA das folhas de todas as cultivares combinadas, separadas de acordo com o tipo sexual do mamoeiro, femininas e hermafroditas. ....46

### **CAPÍTULO 3 - Classificação de cultivares de mamoeiros por espectroscopia no infravermelho próximo combinada à PLS-DA e machine learning**

**Tabela. 1.** Número de sementes e folhas utilizadas no conjunto de calibração (treinamento), validação cruzada (teste) das cultivares de mamoeiro. ....64

**Tabela. 2.** Performance dos modelos de classificação desenvolvidos com os espectros NIR das sementes de cinco cultivares de mamoeiro para os conjuntos calibração (treinamento), validação cruzada (teste) e validação externa.....65

**Tabela. 3.** Performance dos modelos de classificação desenvolvidos com os espectros NIR das folhas dos seedlings de cinco cultivares de mamoeiro para os conjuntos calibração (treinamento), validação cruzada (teste) e validação externa..66

## LISTA DE FIGURAS

### **CAPÍTULO 2 - Determinação sexual em sementes e folhas de mamoeiros utilizando espectroscopia no infravermelho próximo combinadas a técnicas multivariadas e machine learning**

**Figura 1.** Método do método de referência para a identificação do tipo sexual dos mamoeiros. (A) flores femininas (a – ovário grande e em formado arredondado; b – estigma em forma de leque), (B) flores hermafroditas (a – androceu; b – ovário) e (C) flores masculinas (a – androceu; b – ovário rudimentar).....47

**Figura 2.** Espectros NIR brutos médios das sementes de mamoeiros. (A) espectros de todas as cultivares combinadas e espectros das cultivares (B) ‘Calimosa’, (C) ‘Formosa’, (D) ‘T2’, (E) ‘Ouro’ e (F) ‘THB’ separadas de acordo com o tipo sexual do mamoeiro (femininas e hermafroditas).....48

**Figura 3.** Espectros NIR brutos médios das folhas de mamoeiros. (A) espectros de todas as cultivares combinadas e espectros médios das cultivares (B) ‘Calimosa’, (C) ‘Formosa’, (D) ‘T2’ e (E) ‘THB’ separadas de acordo com o tipo sexual do mamoeiro (femininas e hermafroditas).....49

### **CAPÍTULO 3 - Classificação de cultivares de mamoeiros por espectroscopia no infravermelho próximo combinada à PLS-DA e machine learning**

**Figura. 1.** Espectros NIR brutos totais (A) e médios (B) das sementes de mamoeiros. Espectros de todas as cultivares combinadas e médios das cultivares ‘Calimosa’ (C1), ‘Formosa’ (C2), ‘T2’ (C3), ‘Ouro’ (C4) e ‘THB’ (C5). .....67

**Figura. 2.** Espectros NIR brutos totais (A) e médios (B) das folhas de mamoeiros. Espectros de todas as cultivares combinadas e médios das cultivares ‘Calimosa’ (C1), ‘Formosa’ (C2), ‘T2’ (C3), e ‘THB’ (C5). .....68

## CAPÍTULO 1 - Considerações gerais

### 1. Introdução

O Brasil é um dos principais produtores mundiais de frutas tropicais. Dentre as espécies de frutíferas tropicais cultivadas no país pode-se destacar o mamoeiro (*Carica papaya* L.), que apresenta grande importância social e econômica para o agronegócio brasileiro. De acordo com a *Food and Agriculture Organization* (Faostat, 2021a), o Brasil é o terceiro maior produtor mundial de mamão (1.400.00 toneladas métricas), com 10,72% na produção mundial. Esta produção ocorreu em uma área cultivada de 26.526 hectares (Embrapa, 2019). Em 2018, foram exportadas 42.669,06 toneladas, o que gerou uma renda em torno de US\$ 50.061.420,00 (Mapa, 2019).

Em função da grande incidência de doenças, o cultivo do mamoeiro requer renovações frequentes dos pomares (2,5 a 4 anos), demandando constantes investimentos, em especial na aquisição de sementes e mudas (Silva et al., 2016; Oliveira et al., 1994). Estes materiais propagativos são geralmente obtidos via propagação sexuada em função de sua maior eficiência econômica (Boas, 2019; Farias et al., 2009; Simão, 1998). Todavia, o mamoeiro apresenta três classes de plantas, ou seja, monóicas, dióicas e monóclinas (Ming et al., 2007a; Storey, 1953; Hofmeyr, 1941) e o cruzamento entre estes tipos gera sementes com diferentes tipos sexuais (Marin e Gomes, 1986).

As plantas dos diferentes tipos sexuais produzirão frutos com características distintas, ou seja, as plantas hermafroditas irão produzir frutos de formato alongado, piriforme, com pequena cavidade interna e maior valor comercial (Dantas e Castro Neto, 2000). As plantas femininas irão produzir frutos de formato arredondado ou ligeiramente ovalado, cuja cavidade interna é grande, em relação à espessura da polpa, e necessitam da fecundação por pólen de plantas masculinas ou hermafroditas para produzirem frutos (Costa e Pacova, 2003). Assim, as plantas de mamoeiro podem apresentar flores pistiladas ou femininas típicas, flores hermafroditas e flores estaminadas ou masculinas típicas (Marin et al., 1995; Oliveira et al., 1994). Desta forma, visando garantir o plantio de plantas hermafroditas que

produzem frutos com o formato dos frutos exigidos pelo mercado consumidor, ou seja, frutos mais alongados e com polpa carnosa, é necessário o plantio de pelo menos três mudas por cova, o que acarreta um maior custo de produção (Costa e Pacova, 2003; Simão, 1998).

Em função da importância dos tipos sexuais para a cultura do mamoeiro, vários métodos foram desenvolvidos para identificar os tipos sexuais precocemente, na tentativa de diminuir o número de sementes e/ou mudas a serem utilizados. Desta forma, pode-se citar os métodos citológicos (Datta, 1971), por biomarcadores (Jindal e Singh, 1976), de determinação dos compostos fenólicos (Parasnis et al., 1999), e as ferramentas biomoleculares (Parasnis et al., 2000). Apesar da precisão, estes métodos são dispendiosos e requerem laboratórios especializados, dificultando a sua aplicação em condições de cultivo.

Similarmente, as diferentes cultivares de mamoeiro são comumente caracterizadas por meio de descritores morfológicos (Xu et al., 2009), porém este método é considerado limitado para a cultura do mamoeiro em função da baixa diversidade genética desta espécie, o que dificulta a identificação fenotípica das plantas (Santana et al., 2004). Desta forma, métodos mais acurados de diferenciação das cultivares de mamoeiros vem sendo empregados. Por exemplo, o uso de marcadores moleculares, tais como *Random Amplification of Polymorphic DNA* (RAPD) e microssatélites (*Inter Simple Sequence Repeats* – ISSR e *Simple Sequence Repeats* - SSR) foram relatados por Degel Barbosa et al. (2011) e Ming et al. (2007a), respectivamente. Contudo, estes métodos são destrutivos, geram resíduos químicos, demandam tempo e necessitam de laboratórios especializados para sua execução.

Neste sentido, métodos mais simples e eficazes de seleção de mamoeiros de acordo com o tipo sexual e cultivares, nas sementes e/ou mudas, poderiam minimizar os problemas decorrentes da sexagem nos plantios comerciais. A espectroscopia no infravermelho próximo (NIR) por ser uma técnica analítica rápida e não destrutiva poderia ser utilizada para estas finalidades, uma vez que os espectros NIR são fontes de informações para a identificação qualitativa de amostras (Pasquini, 2003).

Desta forma, a espectroscopia NIR foi utilizada com sucesso para a sexagem de diferentes materiais biológicos, tais como larvas do bicho-da-seda (Tao et al., 2019), bezerros e vacas (O'Neill et al., 2017), em outros. Similarmente, a espectroscopia NIR tem sido utilizada para a diferenciação de cultivares de soja, arroz e feijão (Singh et al., 2018), de cultivares de milho doce (Qiu et al., 2019) e nozes de macadâmia (Rahman et al., 2021), porém não foram encontrados trabalhos relacionados à sexagem e à classificação de cultivares de mamoeiros.

Portanto, o objetivo geral deste trabalho foi verificar a possibilidade de se utilizar a espectroscopia NIR como um método não destrutivo para a sexagem de mamoeiros e à classificação de cultivares, por objetivos específicos: *i.* verificar a possibilidade de se utilizar as sementes e folhas dos *seedlings* para sexagem de mamoeiros e *ii.* desenvolver modelos matemáticos visando a predição do tipo sexual dos mamoeiros e à classificação de cultivares.

## 5. Conclusão

É possível a classificação das cultivares de mamoeiro ‘THB’ e ‘Ouro’ utilizando sementes por meio da espectroscopia no infravermelho próximo (NIR) combinada à técnica multivariada de calibração PLS-DA. No entanto, não foi possível classificar corretamente as demais cultivares utilizando tanto as sementes quanto as folhas dos *seedlings*.

Os modelos gerados podem ser utilizados para a certificação da origem do material vegetal, bem como no controle de fraudes. Entretanto, são necessários mais estudos visando melhorar o desempenho dos modelos desenvolvidos, por exemplo com uso de seleção de variáveis.

## 6. Referências

Balabin, R.M., Lomakina, E.I., 2011. Support vector machine regression (SVR/LS-SVM) - An alternative to neural networks (ANN) for analytical chemistry? Comparison of nonlinear methods on near infrared (NIR) spectroscopy data. *Analyst* 136, 1703–1712. <https://doi.org/10.1039/c0an00387e>

Ballabio, D., Consonni, V., 2013. Classification tools in chemistry. Part 1: Linear models. PLS-DA. *Anal. Methods*. <https://doi.org/10.1039/c3ay40582f>

Cortes, C., Vapnik, V., 1995. Support-vector networks. *Mach. Learn.* 20, 273–297. <https://doi.org/10.1007/bf00994018>

Dantas, J.L.L., Junghans, D.T., Lima, J.F. de, 2013. Mamão: o produtor pergunta, a Embrapa responde., 2<sup>a</sup> rev. e atual. ed. Embrapa, Brasília, DF .

Degel Barbosa, C., Pio Viana, A., Silva, S., Quintal, R., Pereira, M.G., 2011. CD Barbosa et al. Brazilian Society of Plant Breeding. Printed in Brazil, Crop Breeding and Applied Biotechnology.

FAOSTAT, F. and A.O. of the U.N.S., 2021. Major tropical fruits - Statistical compendium 2021., in: Statistics Division. Rome, pp. 20–31.

Fernandes, D.D.S., Gomes, A.A., Costa, G.B. Da, Silva, G.W.B.D., Vêras, G., 2011. Determination of biodiesel content in biodiesel/diesel blends using NIR and visible spectroscopy with variable selection. *Talanta* 87, 30–34. <https://doi.org/10.1016/j.talanta.2011.09.025>

Ferreira, M.M.C., 2015. Quimiometria: conceitos, métodos e aplicações, in:

UNICAMP (Ed.), .

Haq, Q.M.I., Mabood, F., Naureen, Z., Al-Harrasi, A., Gilani, S.A., Hussain, J., Jabeen, F., Khan, A., Al-Sabari, R.S.M., Al-khanbashi, F.H.S., Al-Fahdi, A.A.M., Al-Zaabi, A.K.A., Al-Shuraiqi, F.A.M., Al-Bahaisi, I.M., 2018. Application of reflectance spectroscopies (FTIR-ATR & FT-NIR) coupled with multivariate methods for robust in vivo detection of begomovirus infection in papaya leaves. *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* 198, 27–32. <https://doi.org/10.1016/j.saa.2018.02.065>

Kennard, R.W., Stone, L.A., 1969. Computer Aided Design of Experiments. *Technometrics* 11, 137–148. <https://doi.org/10.1080/00401706.1969.10490666>

Kusumaningrum, D., Lee, H., Lohumi, S., Mo, C., Kim, M.S., Cho, B.K., 2018. Non-destructive technique for determining the viability of soybean (*Glycine max*) seeds using FT-NIR spectroscopy. *J. Sci. Food Agric.* 98, 1734–1742. <https://doi.org/10.1002/jsfa.8646>

Lopes Cruz, J., Ferreira Coelho, E., Pelacani, C.R., Coelho Filho, M.A., Tosta Dias, A., Taluana Dos Santos, M., 2004. Growth and dry matter and carbon partition in papaya plants in response to nitrogen nutrition. *Bragantia* 63, 351–361. <https://doi.org/10.1590/s0006-87052004000300005>

Mazumder, R., Friedman, J.H., Hastie, T., 2011. SparseNet: Coordinate descent with nonconvex penalties. *J. Am. Stat. Assoc.* 106, 1125–1138. <https://doi.org/10.1198/jasa.2011.tm09738>

Ming, R., Yu, Q., Moore, P.H., 2007. Sex determination in papaya. *Semin. Cell Dev. Biol.* <https://doi.org/10.1016/j.semcdb.2006.11.013>

Miralbés, C., 2008. Discrimination of European wheat varieties using near infrared reflectance spectroscopy. *Food Chem.* 106, 386–389. <https://doi.org/10.1016/j.foodchem.2007.05.090>

Orrillo, I., Cruz-Tirado, J.P., Cardenas, A., Oruna, M., Carnero, A., Barbin, D.F., Siche, R., 2019. Hyperspectral imaging as a powerful tool for identification of papaya seeds in black pepper. *Food Control* 101, 45–52. <https://doi.org/10.1016/j.foodcont.2019.02.036>

Pasquini, C., 2018. Near infrared spectroscopy: A mature analytical technique with new perspectives – A review. *Anal. Chim. Acta.* <https://doi.org/10.1016/j.aca.2018.04.004>

Pasquini, C., 2003. Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications. *J. Braz. Chem. Soc.* <https://doi.org/10.1590/S0103-50532003000200006>

Paull, R.E., Duarte, O., 2011. Tropical fruits. CABI.

Purcell, D.E., O'shea, M.G., Johnson, R.A., Kokot, S., 2009. Near-infrared spectroscopy for the prediction of disease ratings for fiji leaf gall in sugarcane clones. *Appl. Spectrosc.* 63, 450–457. <https://doi.org/10.1366/000370209787944370>

Qiu, G., Lü, E., Wang, N., Lu, H., Wang, F., Zeng, F., 2019. Cultivar classification of single sweet corn seed using fourier transform near-infrared spectroscopy combined with discriminant analysis. *Appl. Sci.* 9, 1530. <https://doi.org/10.3390/app9081530>

Rahman, A., Wang, S., Yan, J., Xu, H., 2021. Intact macadamia nut quality assessment using near-infrared spectroscopy and multivariate analysis. *J. Food Compos. Anal.* 102, 104033. <https://doi.org/10.1016/j.jfca.2021.104033>

Sandoval, K.V., Vila, D.D., Gracia, T.J.H., 2017. Estudio del mercado de papaya mexicana: un análisis de su competitividad (2001-2015) (Study of the Mexican Papaya Market: An Analysis of Its Competitiveness (2001–2015)). undefined.

Santana, L.R.R., Matsuura, F.C.A.U., Cardoso, R.L., 2004. Genótipos melhorados de mamão (*Carica papaya* L.): avaliação sensorial e físico-química dos frutos. *Ciência e Tecnol. Aliment.* 24, 217–222. <https://doi.org/10.1590/s0101-20612004000200010>

Savitzky, A., Golay, M.J.E., 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* 36, 1627–1639. <https://doi.org/10.1021/ac60214a047>

Singh, S., Patel, S., Litoria, N., Gandhi, K., Faldu, P., Patel, K.G., 2018. Comparative Efficiency of Conventional and NIR Based Technique for Proximate Composition of Pigeon Pea, Soybean and Rice Cultivars. *Int. J. Curr. Microbiol. Appl. Sci.* 7, 773–782. <https://doi.org/10.20546/ijcmas.2018.701.094>

Sjöström, M., Wold, S., Söderström, B., 1986. PLS DISCRIMINANT PLOTS, in: *Pattern Recognition in Practice*. Elsevier, pp. 461–470. <https://doi.org/10.1016/b978-0-444-87877-9.50042-x>

Wong, K.H., Razmovski-Naumovski, V., Li, K.M., Li, G.Q., Chan, K., 2013. Differentiation of *Pueraria lobata* and *Pueraria thomsonii* using partial least square discriminant analysis (PLS-DA). *J. Pharm. Biomed. Anal.* 84, 5–13. <https://doi.org/10.1016/j.jpba.2013.05.040>

Xu, H.R., Yu, P., Fu, X.P., Ying, Y. Bin, 2009. On-site variety discrimination of tomato plant using visible-near infrared reflectance spectroscopy. *J. Zhejiang Univ. Sci. B* 10, 126–132. <https://doi.org/10.1631/jzus.B0820200>