

RESSALVA

Atendendo solicitação do(a) autor(a), o texto completo será disponibilizado somente a partir de 29/01/2027

At the author's request, the full text will not be available online until January 29, 2027

FELIPE SOUZA CARVALHO

**INOVAÇÕES METODOLÓGICAS NA AVALIAÇÃO DA QUALIDADE DE
SEMENTES: MODELAGEM MATEMÁTICA E IMAGEM MULTIESPECTRAL
ASSOCIADA AO APRENDIZADO DE MÁQUINA**

Botucatu

2026

FELIPE SOUZA CARVALHO

**INOVAÇÕES METODOLÓGICAS NA AVALIAÇÃO DA QUALIDADE DE
SEMENTES: MODELAGEM MATEMÁTICA E IMAGEM MULTIESPECTRAL
ASSOCIADA AO APRENDIZADO DE MÁQUINA**

Tese apresentada à Faculdade de
Ciências Agronômicas da Unesp Câmpus
de Botucatu, para obtenção do título de
Doutor em Agronomia (Agricultura).

Orientadora: Profa. Dra. Maria Márcia
Pereira Sartori

Botucatu
2026

C331i

Carvalho, Felipe Souza

INOVAÇÕES METODOLÓGICAS NA AVALIAÇÃO DA
QUALIDADE DE SEMENTES: MODELAGEM MATEMÁTICA E
IMAGEM MULTIESPECTRAL ASSOCIADA AO
APRENDIZADO DE MÁQUINA / Felipe Souza Carvalho. --
Botucatu, 2026
70 p.

Tese (doutorado) - Universidade Estadual Paulista (UNESP),
Faculdade de Ciências Agrônômicas, Botucatu
Orientadora: Maria Márcia Pereira Sartori

1. Sementes. I. Título.

CERTIFICADO DE APROVAÇÃO

Título:

INOVAÇÕES METODOLÓGICAS NA AVALIAÇÃO DA QUALIDADE DE SEMENTES: MODELAGEM MATEMÁTICA E IMAGEM MULTIESPECTRAL ASSOCIADA AO APRENDIZADO DE MÁQUINA

AUTOR: FELIPE SOUZA CARVALHO

ORIENTADORA: MARIA MÁRCIA PEREIRA SARTORI

Aprovado como parte das exigências para obtenção do Título de Doutor em Agronomia (Agricultura), pela Comissão Examinadora:


Pesquisadora Dr.^a MARIA MÁRCIA PEREIRA SARTORI (Participação Presencial)
Departamento de Produção Vegetal / UNESP / Câmpus de Botucatu - FCA


Prof. Dr. RENATO FERNANDES CANTÃO (Participação Virtual)
Departamento de Física, Química e Matemática / Universidade Federal de São Carlos


Dr. THIAGO BARBOSA BATISTA (Participação Presencial)
Centro de Energia Nuclear na Agricultura / Universidade de São Paulo


Prof. Dr. EDVALDO APARECIDO AMARAL DA SILVA (Participação Presencial)
Departamento de Produção Vegetal / UNESP / Câmpus de Botucatu - FCA


Dr.^a BRUNNA RITHIELLY REZENDE (Participação Presencial)
Botucatu/SP

Botucatu, 29 de janeiro de 2026.

*A minha amada família,
Solange, Solonildo e Karoline,
dedico*

AGRADECIMENTOS

Escrever os agradecimentos foi, sinceramente, a parte que eu mais esperei. Confesso que romantizei bastante esse momento na minha cabeça, mas quando finalmente me sentei para escrever, me veio o clichê de querer citar uma infinidade de nomes. Para me ajudar, fui buscar referências nos agradecimentos de colegas e, em um deles, encontrei a seguinte frase: *“But a wise person told me I should write with my heart and let my personality be part of that, especially because, according to him, the acknowledgments are the most read part of a thesis.”* Não sei quem disse isso, mas fui tocado. Então, *let’s write with the heart*.

É doido pensar na minha trajetória até aqui. Às vezes, custo a acreditar o quão longe cheguei e na sorte imensa que tenho de ter tantas pessoas especiais na minha vida. Gostaria de começar agradecendo à minha família.

À minha mãe, por simplesmente ser a melhor mãe que eu poderia ter. Eu te amo profundamente. Ao meu pai, por tudo, absolutamente tudo. O amo muito! E à minha querida *hermana*, te amo também!

Sem vocês, eu simplesmente não seria ninguém.

Agradeço à Ufra, instituição que amo, e que me proporcionou oportunidades que transformaram completamente a minha vida. Em especial à professora Luciana Borges, que foi quem plantou a sementinha que um dia me fez pensar em seguir na pós-graduação e que, mais do que plantar, me ajudou a cultivar esse sonho desde o início e, principalmente, verdadeiramente acreditou em mim.

Outra pessoa muito querida, de extrema importância nessa jornada, foi a Mi Lima. Desde o primeiro contato, quando eu ainda buscava informações sobre a pós, até literalmente me abrigar em sua casa nas minhas primeiras semanas em Botucatu, sem nem me conhecer direito. Depois, tornou-se uma grande amiga, e ainda me apresentou à sua Clyde (‘vulgo’ Brunna), com quem também construí uma amizade linda. Juntas, tornaram essa caminhada, que por vezes pesada, muito mais leve.

Aproveito para agradecer também às minhas queridas amigas do grupo de pesquisa ModelStat: Mônica, Gabriela e Rêzinha, obrigada por cada conversa, cada risada e cada aprendizado compartilhado.

Na mesma linha, deixo minha eterna gratidão à professora Maria Márcia Pereira Sartori, minha orientadora. Muito mais que orientadora, um apoio, um porto seguro em Botucatu. Obrigado por me fazer enxergar a estatística com outros olhos. Aprendi

tanto com a senhora... Obrigado por me acolher como um filho, por proporcionar momentos leves, descontraídos e inesquecíveis, e, acima de tudo, por acreditar em mim.

Agradeço também ao professor Amaral, a quem tenho profunda admiração pelo trabalho. Obrigado pelos ensinamentos e pelo apoio durante esses anos, e por me fazer ainda mais apaixonado pelo universo das sementes.

Aos colegas do Laboratório de Sementes: Gustavo, Dennis, Ana, Vitor, Taise, Tamires, as queridas que já citei e à nossa chefe Val: saibam que ter vocês como colegas e amigos é um privilégio que poucos têm. A jornada ao lado de vocês foi leve e prazerosa.

Agradeço ainda aos colegas, professores e técnicos do melhor departamento da Unesp — o Departamento de Melhoramento e Produção Vegetal.

E a todas as pessoas que cruzaram meu caminho durante os anos em Botucatu — amigos da república Monte Olimpo e Saia de Baixo, que me acolheram de braços abertos e me fizeram sentir em casa.

Finalizo agradecendo à própria Botucatu, essa cidade maravilhosa onde vivi alguns dos melhores anos da minha vida. Que desafio foi largar tudo em Paragominas, no interior do Pará, e me aventurar em uma cidade cujo nome eu mal conhecia. Mas, olhando para trás, que bom que tive essa coragem — ela me transformou em uma pessoa melhor.

Obrigado, Botucatu e Unesp, por esse privilégio de carregar comigo a honra de ser um Unespiano, e com muito orgulho!

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) – Código de Financiamento 001.

RESUMO

A indústria de sementes tem investido continuamente em estratégias voltadas à garantia de alta qualidade, com o propósito de maximizar o desempenho das sementes em campo, reduzir os riscos associados ao armazenamento e fortalecer a competitividade ao longo da cadeia produtiva. Contudo, os métodos tradicionais de avaliação da qualidade de sementes permanecem, em grande parte, demorados, subjetivos e destrutivos, o que tem impulsionado a busca por alternativas mais precisas, objetivas e eficientes. Nesse cenário, abordagens baseadas em modelagem matemática e em técnicas de imageamento multiespectral surgem como estratégias complementares e inovadoras, capazes de aprimorar a compreensão e o monitoramento da qualidade das sementes.

A modelagem matemática permite extrair informações preditivas a partir de dados experimentais obtidos por métodos convencionais, otimizando o aproveitamento dos resultados e favorecendo estimativas mais robustas da longevidade. Com base nessa premissa, o primeiro estudo (*Modeling Seed Longevity and Percentile Prediction: A Sigmoidal Function Approach in Soybean, Maize, and Tomato*) investigou a longevidade de sementes de soja, milho e tomate armazenadas sob condições controladas, utilizando os modelos sigmóides Logístico e de Boltzmann. O modelo Logístico apresentou melhor desempenho na estimativa dos percentis de viabilidade (P50, P85 e P25), com acurácia de 83%, enquanto o modelo de Boltzmann alcançou 54%. Os resultados evidenciaram diferenças entre cultivares quanto à taxa e ao padrão de deterioração, destacando o potencial da modelagem matemática como ferramenta preditiva para otimização do armazenamento, redução de perdas econômicas e apoio a programas de conservação genética.

O segundo estudo (*A Multispectral Imaging Framework for Discriminating Primed and Unprimed Seeds Using Machine Learning Algorithms*) avaliou o uso da imagem multiespectral (MSI) associada a algoritmos de aprendizado de máquina para discriminar sementes de *Solanum lycocarpum* submetidas ao *priming* de sementes. Os algoritmos Random Forest e Redes Neurais apresentaram os melhores resultados, com acurácia de até 89,8%, demonstrando a sensibilidade da MSI em detectar alterações fisiológicas sutis induzidas pelo *priming*, de forma rápida e não destrutiva. De forma integrada, os resultados desta tese evidenciam o potencial da combinação entre modelagem matemática e técnicas de imageamento multiespectral como abordagens complementares para a avaliação da qualidade de sementes. Essas

estratégias oferecem bases científicas e tecnológicas para aprimorar o controle de qualidade, otimizar o armazenamento e a rastreabilidade, além de promover avanços significativos na conservação e no manejo de sementes, tanto em contextos comerciais quanto de pesquisa.

Palavras-chave: Qualidade de sementes; Modelagem matemática; Imagem multiespectral; Aprendizado de máquina.

ABSTRACT

The seed industry has continuously invested in strategies aimed at ensuring high quality, with the purpose of maximizing field performance, reducing risks associated with storage, and strengthening competitiveness throughout the production chain. However, traditional methods for assessing seed quality remain largely time-consuming, subjective, and destructive, which has driven the search for more precise, objective, and efficient alternatives. In this context, approaches based on mathematical modeling and multispectral imaging techniques have emerged as complementary and innovative strategies capable of improving the understanding and monitoring of seed quality.

Mathematical modeling allows the extraction of predictive information from experimental data obtained through conventional methods, optimizing the use of results and providing more robust estimates of seed longevity. Based on this premise, the first study (*Modeling Seed Longevity and Percentile Prediction: A Sigmoidal Function Approach in Soybean, Maize, and Tomato*) investigated the longevity of soybean, maize, and tomato seeds stored under controlled conditions, using Logistic and Boltzmann sigmoidal models. The Logistic model showed superior performance in estimating viability percentiles (P50, P85, and P25), achieving an accuracy of 83%, while the Boltzmann model reached 54%. The results revealed differences among cultivars regarding deterioration rates and patterns, highlighting the potential of mathematical modeling as a predictive tool for optimizing storage, reducing economic losses, and supporting genetic conservation programs.

The second study (*A Multispectral Imaging Framework for Discriminating Primed and Unprimed Seeds Using Machine Learning Algorithms*) evaluated the use of multispectral imaging (MSI) combined with machine learning algorithms to discriminate *Solanum lycocarpum* seeds subjected to seed priming. The Random Forest and Neural Network algorithms achieved the best results, with accuracies of up to 89.8%, demonstrating the sensitivity of MSI in detecting subtle physiological changes induced by priming in a fast and non-destructive manner.

Taken together, the results of this thesis demonstrate the potential of combining mathematical modeling and multispectral imaging techniques as complementary approaches for assessing seed quality. These strategies provide scientific and technological foundations to enhance quality control, optimize storage and traceability, and promote significant advances in seed conservation and management, both in commercial and research contexts.

Keywords: Seed quality; Mathematical modeling; Multispectral imaging; Machine learning.

SUMÁRIO

INTRODUÇÃO GERAL.....	15
CHAPTER 1 – MODELING SEED LONGEVITY AND PERCENTILE PREDICTION: A SIGMOIDAL FUNCTION APPROACH IN SOYBEAN, MAIZE, AND TOMATO.....	17
1.1 INTRODUCTION.....	18
1.2 MATERIALS AND METHODS.....	20
1.2.1 Biological Material.....	20
1.2.2 Water Content.....	21
1.2.3 Germination.....	21
1.2.4 Longevity.....	21
1.2.5 Statistical Methodology.....	21
1.2.6 Fitting Longevity Predictors.....	23
1.2.7 Percentile Estimation.....	23
1.3 RESULTS.....	24
1.4 DISCUSSION.....	37
1.5 CONCLUSION.....	40
REFERENCES.....	41
CAPÍTULO 2 – A MULTISPECTRAL IMAGING FRAMEWORK FOR DISCRIMINATING PRIMED AND UNPRIMED SEEDS USING MACHINE LEARNING ALGORITHMS.....	45
2.1 INTRODUCTION.....	46
2.2 MATERIALS AND METHODS.....	47
2.2.1 Biological Material.....	47
2.2.2 Hydropriming.....	48
2.2.3 Moisture content.....	48
2.2.4 Physiological quality characterization.....	49
2.2.5 Multispectral Image Acquisition.....	49
2.2.6 Data Analysis.....	50
2.3 RESULTS.....	51
2.3.1 Seed characterization.....	51
2.3.2 Multispectral image markers in the segmentation of hydropriming and	

	untreated control seeds.....	52
2.3.3	Performance of Discriminant Analysis and Machine Learning Models in Classifying Hydropriming and Untreated Control Seeds.....	54
2.4	DISCUSSION.....	58
2.5	CONCLUSION.....	61
	REFERENCES.....	61
	CONSIDERAÇÕES FINAIS.....	64
	REFERÊNCIAS.....	65

INTRODUÇÃO GERAL

A qualidade das sementes é um atributo multifatorial, resultante da interação de características físicas, fisiológicas, sanitárias e genéticas, as quais determinam o desempenho das sementes (ElMasry, Mandour, Al-Rejaie, *et al.*, 2019). A indústria de sementes tem investido continuamente em estratégias que assegurem alta qualidade, com o objetivo de maximizar o desempenho das sementes no campo, reduzir riscos durante o armazenamento e aumentar a competitividade da cadeia produtiva (ElMasry, Mandour, Al-Rejaie, *et al.*, 2019). Nesse contexto, um dos focos crescentes da indústria é o aprimoramento do controle de qualidade interno que monitora a qualidade das sementes durante toda a cadeia de produção. No entanto, os métodos tradicionais para avaliação da qualidade das sementes apresentam limitações significativas (Baek *et al.*, 2019). Esses métodos, que incluem inspeções físicas e químicas e análises visuais, são frequentemente caros, demorados e, principalmente, subjetivos, pois dependem da interpretação do analista. Além disso, muitos desses testes possuem caráter destrutivo, inviabilizando a reutilização das amostras (Baek *et al.*, 2019; Boelt *et al.*, 2018).

Diante dessas limitações, cresce o interesse por métodos inovadores capazes de reduzir a subjetividade, otimizar o tempo de análise e fornecer resultados mais precisos, consistentes e preditivos (ElMasry, Mandour, Wagner, *et al.*, 2019). Nesse contexto, técnicas baseadas em modelagem matemática, imagem multiespectral (MSI) e a espectrometria no infravermelho próximo (NIRS) têm ganhado destaque como alternativas promissoras para o monitoramento da qualidade de sementes por serem métodos rápidos, não destrutivos e que não necessitam de pré-tratamentos (Chu *et al.*, 2024; ElMasry, Mandour, Wagner, *et al.*, 2019; Fonseca De Oliveira *et al.*, 2022; Liu *et al.*, 2019; Mortensen *et al.*, 2021).

Modelos matemáticos para avaliar a longevidade de sementes são ferramentas essenciais tanto para o manejo de bancos de germoplasma quanto para a gestão comercial de sementes. Estimativas precisas da taxa de perda de viabilidade de lotes armazenados fornecem informações valiosas que subsidiam decisões estratégicas de manejo (De Faria *et al.*, 2020). Essas informações permitem preservar ao máximo a viabilidade das sementes no período pós-colheita até a comercialização, direcionando, por exemplo, lotes com maior viabilidade para demandas específicas e estratégicas (Hay *et al.*, 2018; Lee *et al.*, 2019).

No contexto da preservação de recursos genéticos em bancos de germoplasma, a estimativa da longevidade com o uso de modelos matemáticos também desempenha um papel crucial na otimização do processo de renovação dos acessos (Genebank standards for plant genetic resources for food and agriculture, 2016; Hay *et al.*, 2018). Essa abordagem contribui para evitar a erosão genética e a perda excessiva de viabilidade, além de reduzir o consumo desnecessário de amostras em testes de monitoramento. Como resultado, promove

um manejo mais eficiente, tanto em termos de recursos materiais quanto de mão de obra, assegurando a conservação dos recursos genéticos a longo prazo.

A imagem multiespectral de sementes é uma técnica não destrutiva que permite a obtenção simultânea de informações espaciais e espectrais das sementes por meio da captura de imagens de reflectância de superfície. Essa técnica possibilita a análise da química de superfície e de atributos morfológicos, como cor, textura e forma. As imagens multiespectrais são consideradas um intermediário entre as imagens RGB convencionais e as imagens hiperespectrais, capturando geralmente de 3 a 20 bandas ou planos de imagem, em vez de faixas contínuas (Mortensen *et al.*, 2021; Su e Sun, 2018).

Essas imagens fornecem registros detalhados das variações espectrais em amostras heterogêneas, possibilitando a aplicação de métodos estatísticos para identificar diferenças entre e dentro das amostras. Isso permite detectar padrões subjacentes que, de outra forma, não seriam evidentes a olho nu ou por técnicas tradicionais (ElMasry, Mandour, Al-Rejaie, *et al.*, 2019).

CONSIDERAÇÕES FINAIS

Os resultados obtidos nesta tese evidenciam o potencial de métodos inovadores na avaliação da qualidade de sementes, demonstrando que abordagens quantitativas e não destrutivas podem ampliar significativamente a compreensão sobre o comportamento fisiológico das sementes e a eficiência dos processos de controle de qualidade.

A modelagem matemática baseada em funções sigmóides mostrou-se uma ferramenta robusta e versátil para a estimativa de percentis de longevidade (P85, P50 e P25), capturando com precisão os padrões naturais de deterioração em espécies de importância agrícola, como soja, milho e tomate. A aplicação do modelo Logístico apresentou desempenho superior, com boa capacidade de generalização entre culturas, o que reforça seu potencial para uso em contextos comerciais e de conservação. Ainda assim, reconhece-se a necessidade de ampliar o número de lotes e espécies avaliadas, especialmente para o tomate, a fim de refinar as estimativas e consolidar a aplicabilidade dos modelos em diferentes cenários de armazenamento e qualidade inicial.

A técnica de imagem multiespectral, aliada a algoritmos de aprendizado de máquina, também se mostrou altamente promissora, permitindo detectar modificações fisiológicas sutis decorrentes do *priming* de sementes de *Solanum lycocarpum*. Essa abordagem não destrutiva demonstrou elevado poder discriminativo, viabilizando análises rápidas, objetivas e automatizáveis, com potencial para aplicação prática em programas de monitoramento, rastreabilidade e controle de qualidade de sementes. De forma integrada, os dois estudos reforçam a importância da incorporação de ferramentas matemáticas e computacionais avançadas no contexto da tecnologia de sementes. A combinação entre modelagem preditiva e imageamento multiespectral representa um avanço metodológico relevante, capaz de reduzir a subjetividade, otimizar o tempo de análise e fornecer informações mais consistentes para subsidiar decisões estratégicas na produção, armazenamento e conservação de sementes.

REFERÊNCIAS

- BAEK, I.; KUSUMANINGRUM, D.; KANDPAL, L. M.; LOHUMI, S.; MO, C.; KIM, M. S.; CHO, B.-K. Rapid Measurement of Soybean Seed Viability Using Kernel-Based Multispectral Image Analysis. **Sensors**, v. 19, n. 2, p. 271, 11 jan. 2019.
- BOELT, B.; SHRESTHA, S.; SALIMI, Z.; JØRGENSEN, J. R.; NICOLAISEN, M.; CARSTENSEN, J. M. Multispectral imaging – a new tool in seed quality assessment? **Seed Science Research**, v. 28, n. 3, p. 222–228, set. 2018.
- CHU, Y. R.; JO, M. S.; KIM, G. E.; PARK, C. H.; LEE, D. J.; CHE, S. H.; NA, C. S. Non-Destructive Seed Viability Assessment via Multispectral Imaging and Stacking Ensemble Learning. **Agriculture**, v. 14, n. 10, p. 1679, 26 set. 2024.
- DE FARIA, R. Q.; DOS SANTOS, A. R. P.; AMORIM, D. J.; CANTÃO, R. F.; DA SILVA, E. A. A.; SARTORI, M. M. P. Probit or Logit? Which is the better model to predict the longevity of seeds? **Seed Science Research**, v. 30, n. 1, p. 49–58, mar. 2020.
- ELMASRY, G.; MANDOUR, N.; AL-REJAIE, S.; BELIN, E.; ROUSSEAU, D. Recent Applications of Multispectral Imaging in Seed Phenotyping and Quality Monitoring—An Overview. **Sensors**, v. 19, n. 5, p. 1090, 4 mar. 2019.
- ELMASRY, G.; MANDOUR, N.; WAGNER, M.-H.; DEMILLY, D.; VERDIER, J.; BELIN, E.; ROUSSEAU, D. Utilization of computer vision and multispectral imaging techniques for classification of cowpea (*Vigna unguiculata*) seeds. **Plant Methods**, v. 15, n. 1, p. 24, dez. 2019.
- FONSECA DE OLIVEIRA, G. R.; MASTRANGELO, C. B.; HIRAI, W. Y.; BATISTA, T. B.; SUDKI, J. M.; PETRONILIO, A. C. P.; CRUSCIOL, C. A. C.; AMARAL DA SILVA, E. A. An Approach Using Emerging Optical Technologies and Artificial Intelligence Brings New Markers to Evaluate Peanut Seed Quality. **Frontiers in Plant Science**, v. 13, p. 849986, 14 abr. 2022.
- Genebank standards for plant genetic resources for food and agriculture.** Revised edition ed. Rome: Food and Agriculture Organization of the United Nations, 2016.
- HAY, F. R.; VALDEZ, R.; LEE, J.-S.; STA. CRUZ, P. C. Seed longevity phenotyping: recommendations on research methodology. **Journal of Experimental Botany**, 16 out. 2018.
- LEE, J.-S.; VELASCO-PUNZALAN, M.; PACLEB, M.; VALDEZ, R.; KRETZSCHMAR, T.; MCNALLY, K. L.; ISMAIL, A. M.; CRUZ, P. C. S.; SACKVILLE HAMILTON, N. R.; HAY, F. R. Variation in seed longevity among diverse Indica rice varieties. **Annals of Botany**, v. 124, n. 3, p. 447–460, 18 out. 2019.
- LIU, W.; XU, X.; LIU, C.; ZHENG, L. Rapid Discrimination of High-Quality Watermelon Seeds by Multispectral Imaging Combined with Chemometric Methods. **Journal of Applied Spectroscopy**, v. 85, n. 6, p. 1044–1049, jan. 2019.

MORTENSEN, A. K.; GISLUM, R.; JØRGENSEN, J. R.; BOELT, B. The Use of Multispectral Imaging and Single Seed and Bulk Near-Infrared Spectroscopy to Characterize Seed Covering Structures: Methods and Applications in Seed Testing and Research. **Agriculture**, v. 11, n. 4, p. 301, 1 abr. 2021.

SU, W.; SUN, D. Multispectral Imaging for Plant Food Quality Analysis and Visualization. **Comprehensive Reviews in Food Science and Food Safety**, v. 17, n. 1, p. 220–239, jan. 2018.