Universidade de São Paulo
Escola Superior de Agricultura "Luiz de Queiroz"

Desenvolvimento de ferramentas de geotecnologia para análise da qualidade da lavoura e estimativa da taxa de germinação aplicadas à cultura da soja

Michaela Bárbara Neto

Dissertação apresentada para obtenção do título de Mestra em Ciências. Área de concentração: Engenharia de Sistemas Agrícolas

Piracicaba
2022

Michaela Bárbara Neto
Engenheira Hídrica

Desenvolvimento de ferramentas de geotecnologia para análise da qualidade da lavoura e estimativa da taxa de germinação aplicadas à cultura da soja

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientador:
Prof. Dr. **PETERSON RICARDO FIORIO**

Dissertação apresentada para obtenção do título de Mestra em Ciências. Área de concentração: Engenharia de Sistemas Agrícolas

Piracicaba
2022

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP

Neto, Michaela Bárbara

Desenvolvimento de ferramentas de geotecnologias para análise da qualidade da lavoura e estimativa da taxa de germinação aplicadas à cultura da soja / Michaela Bárbara Neto. - - versão revisada de acordo com a resolução CoPGR 6018 de 2011 - - Piracicaba, 2022.

56 p.

Dissertação (Mestrado) - - USP / Escola Superior de Agricultura "Luiz de Queiroz".

1. RESAP 2. Posição do Red-Edge 3. ANNSEED 4. Rede neural artificial 5. Integração de dados I. Título

DEDICATÓRIA

In memoriam de José Augusto de Moraes Preto,
que sua alegria, resiliência e
maneira leve de encarar a vida
sempre me acompanhem.

Dedico a você este trabalho.

AGRADECIMENTOS

Agradeço a Deus por me conceder saúde, por permitir que eu esteja hoje com vocês, por estar comigo nos momentos mais difíceis e, por iluminar meus pensamentos me guiando na minha jornada.

À minha família meus pais, João e Solange, e a minha irmã, Dominicky, por toda renúncia e apoio que sempre deram para os meus estudos. Vocês nunca mediram esforços e são meus maiores incentivadores nessa caminhada.

À família Maciel e Fagundes, por acompanharem de perto toda minha trajetória e fazerem com que os dias fossem mais leve. Em especial ao Augusto por todo carinho e incentivo ao longo desses anos, sempre acreditando na minha capacidade de alçar voos mais altos.

À equipe do Water Institute da University of South Florida em especial ao Dr. Shawn Landry e Msc. David Eilers vocês me acolheram durante a minha graduação e me apresentaram o mundo das Geotecnologias, sem a generosidade de vocês essa jornada jamais teria acontecido.

À Escola Superior de Agricultura "Luiz de Queiroz" – ESALQ, ao Programa de Pós-Graduação em Engenharia de Sistemas Agrícolas e ao Departamento de Engenharia de Biossistemas, por promover meu crescimento professional e acadêmico.

Ao meu orientador e amigo, professor Dr. Peterson Ricardo Fiorio, pela paciência, dedicação, ensinamentos, confiança e orientação concedida.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), pela aquisição do equipamento FieldSpec® 3 Jr através do projeto de pesquisa nº 2013/22435-9 e, pela disponibilidade do multiusuários SeedReporter através do projeto de pesquisa nº 2018/03793-5.

À Agência Paulista de Tecnologia dos Agronegócios (APTA), em especial ao Dr. Fábio Luís Ferreira Dias por toda generosidade em transmitir conhecimento, auxílio e por permitir que o experimento fosse realizado.

Aos colaboradores do Laboratório de Análise de Sementes, do Departamento de Produção Vegetal da USP/ESALQ, em especial a Técnica Especializada Superior Me. Helena Pescarin Chamma por toda generosidade em transmitir seus conhecimentos sobre sementes.

Ao Grupo de Pesquisa e Extensão em Topografia e Geoprocessamento – TopoGeo, pelo apoio incondicional no trabalho de campo, em especial Luiz Augusto Leite, Larissa Cardoso, Artur Diaz e Jacqueline Mendonça sem vocês este presente estudo não seria realizado.

Aos professores Dr. Francisco Guilhen Gomes Junior, Dra.Gizele Ingrid Gadotti, Dr.Rodnei Rizzo, Dr.Paulo César Sentelhas, Dr.Leandro Maria Gimenez, pelas valiosas sugestões, ensinamentos e contribuições fundamentais para a conclusão desta.

Aos companheiros de jornada Luiz, Fabiano, Lukas, Thais, Flaviani, Rodolfo, Carlos e Bruna pela amizade, incentivo, apoio e cumplicidade durante essa trajetória.

Enfim, a todos que contribuíram direta ou indiretamente,
Muito obrigada.

EPÍGRAFE

"We can only see a short distance ahead,
but we can see plenty ther
that needs to be done."
Alan Turning

SUMÁRIO

RESUMO **8**

ABSTRACT **9**

1. INTRODUÇÃO **10**

REFERÊNCIAS **11**

2. RESAP: PROPOSTA DE MENSURAÇÃO DA INTENSIDADE DO "RED-EDGE POSITION" E SUA CORRELAÇÃO COM COMPONENTES DE PRODUÇÃO **14**

RESUMO **14**

ABSTRACT **14**

2.1 INTRODUÇÃO **15**

2.2 MATERIAL E MÉTODOS **16**

2.2.1 ÁREA DE ESTUDO 16

2.2.2 EXPERIMENTO 17

2.2.3 CARACTERÍSTICAS AGRONÔMICAS DAS CULTIVARES 17

2.2.4 COLETA DE DADOS IN SITU E EM LABORATÓRIO 18

2.2.4.1 IN SITU 18

2.2.4.2 LABORATÓRIO 20

2.2.5 OBTENÇÃO DA INTENSIDADE DO REP 20

2.2.6 ANÁLISE DE DADOS 22

2.3 RESULTADOS E DISCUSSÃO **23**

2.4 CONCLUSÕES **28**

AGRADECIMENTOS **28**

REFERÊNCIAS **29**

APÊNDICES **33**

3. ANNSEED 1.0: PROPOSTA DE TESTE RÁPIDO DE GERMINAÇÃO **36**

RESUMO **36**

ABSTRACT **36**

3.1 INTRODUÇÃO **37**

3.2 MATERIAL E MÉTODOS **38**

3.2.1 ETAPA 1: AQUISIÇÃO DE DADOS 39

3.2.2 TESTE DE GERMINAÇÃO 39

3.2.3 ETAPA 2: EXTRAÇÃO DAS CARACTERÍSTICAS FÍSICAS E ESPECTRAIS 41

3.2.4 ETAPA 3: ELABORAÇÃO DO ALGORITMO DE PREDIÇÃO 43

3.2.5 TREINO E VALIDAÇÃO 44
3.2.6 REDE NEURAL ARTIFICIAL MULTILAYER PERCEPTRON (RNA-MLP) 44
3.3 RESULTADOS E DISCUSSÃO 46
3.4 CONCLUSÕES 53
3.5 CONSIDERAÇÕES FINAIS 53
AGRADECIMENTOS 53
REFERÊNCIAS 54

RESUMO

Desenvolvimento de ferramentas de geotecnologias para análise da qualidade da lavoura e estimativa da taxa de germinação aplicadas à cultura da soja

Atualmente no cenário mundial o Brasil tem se destacado como o 3º maior exportador de produtos agrícolas do mundo, sendo possivelmente o principal fornecedor das commodities agrícolas na próxima década, com destaque para o complexo soja bastante consumido pelo mercado Asiático. Por essa razão, o uso de novas tecnologias que possibilitem um manejo mais preciso e por consequência um aumento da produção potencial das áreas, tem se tornado um diferencial para os produtores de soja. Diante do exposto, o presente estudo teve como objetivo o desenvolvimento de ferramentas que auxiliem os produtores durante o ciclo da cultura no campo e no pós-colheita, correlacionando os resultados com as componentes de produção e com a taxa de germinação dos lotes. A primeira ferramenta desenvolvida foi fundamentada através da combinação de métodos já consagrados de mensuração da posição do Red-Edge com artifícios trigonométricos, a obtenção de dados espectrais das folhas ocorreu por meio do uso do sensor hiperespectral FieldSpec. Já a segunda ferramenta desenvolvida foi fundamentada essencialmente no aprendizado de máquina, utilizando etapa de visão computacional, espectroscopia e aprendizado em conjunto, a obtenção de dados físicos e espectrais das sementes ocorreu por meio do uso do sensor multiespectral SeedReporter. Os resultados mostraram que durante os estágios fenológicos R3 e R4 a ferramenta fundamentada na zona do Red-Edge exibiu uma correlação positiva muito forte com as taxas de germinação do futuro lote a ser colhido nas semanas seguintes. Já a ferramenta fundamentada em aprendizado de máquina foi capaz de predizer as taxas de germinação dos lotes colhidos em alguns minutos com uma assertividade entorno de 95% para a classe de sementes classificadas como plântulas normais. No geral, a utilização de ambas as ferramentas apresentou grande potencial e uma alternativa viável para os produtores de soja.

Palavras-chave: RESAP; Posição do *Red-Edge*; ANNSEED; Rede neural artificial; Integração de dados

ABSTRACT

Development geotechnologies tools to analyze crops quality and estimate germination rate applied to soybeans crop

Actually, in the world scenario, Brazil has been highlighted as the third biggest exporter of agricultural products, also in the next decade, it may become the main provider, with emphasis on soybean complex which is mainly consumed by the Asian market. Therefore, the use of new technologies that could improve the management of the crops and consequently increases the potential production of areas has become a differential for soybeans producers. The present study aimed to developed tools to assist producers during crop cycles and post-harvest, correlating the results with production components and germination rates of packs. The first tool was based on the combination of already established measured methods of Red-Edge position with trigonometrical artifices, the data was obtained by spectral of leaves collected by FieldSpec hyperspectral sensor. The second tool was essentially based on machine learning by joining computational view, spectroscopy, and learning, the physical and spectral data were obtained from seeds collected by the SeedReporter multispectral sensor. The results demonstrated that during the phenological stages R4 and R5 the tool based on the Red-Edge zone shown a strong positive correlation with germination rates of the future pack to be harvested in the next weeks. The tool based on machine learning was capable to predict the germination rates of harvested packs in few minutes with an accuracy of 95% for seeds classified at the class of normal seedlings. Overall, the utilization of both tools show a high potential and a viable alternative for soybeans producers.

Keywords: RESAP; Red-Edge position; ANNSEED; Artificial neural network; Data integration

1. INTRODUÇÃO

No ano de 2050, segundo as projeções feitas em 2019 pela Organização das Nações Unidas o mundo alcançará a marca de 9,7 bilhões de habitantes destacando as regiões da África Subsaariana e Ásia (ONU, 2019). Esse crescente populacional terá como consequência um aumento significativo da demanda por alimentos (SAATH; FACHINELLO, 2018; NEVES et al., 2020). Na próxima década de acordo com a OECD/FAO (2020), o consumo das principais *commodities* agrícolas deve aumentar expressivamente, com destaque para a demanda por cultivos base para a alimentação humana e produção de ração. No atual cenário, o Brasil se destaca como o 3º maior exportador de produtos agrícolas do mundo, sendo o seu carro chefe o complexo soja (grãos, farelo e óleo) (NEVES et al., 2020).

Ao longo de décadas, a soja [Glycine max (L.) Merril.] tem sido a mais importante *commodity* negociada do Brasil (COELHO et al., 2018). Com uma produção de 135,4 milhões de toneladas na safra 2020/21, o Brasil tem se estabelecido como o maior produtor e exportador da leguminosa no mundo (CONAB, 2021). Atualmente, o cultivo da leguminosa tem se estendido por quase todo território nacional, ocupando cerca de 36 milhões de hectares (APROSOJA, 2019). Essa vasta distribuição pelo território brasileiro, faz com que a produção transcorra nas mais diversas condições ambientais. A cultura em questão possui necessidades climáticas bastante particulares para atingir altos rendimentos, neste contexto, destaca-se a necessidade hídrica e térmica (MUNDSTOCK; THOMAS, 2005; FARIAS; NEPOMUCENO; NEUMAIER, 2007; MORANDO et al., 2014). Em suma, as características desenvolvidas em cada estágio da cultura da soja sofrem influência de diversos fatores (data de plantio, localização geográfica, clima, incidência de pragas, tipo de solo). Além disso, a cultura se subdivide em dois grupos classificados de acordo com o seu crescimento (determinado e indeterminado), sendo estes impactados diretamente pelos fatores expostos anteriormente (FEHR; CAVINESS, 1977; FARIAS; NEPOMUCENO; NEUMAIER, 2007).

Dessa forma, visando promover uma maior eficiência agrícola frente as condições ambientais diversas presentes no território Brasileiro, o uso de novas tecnologias para o monitoramento e controle das lavouras tem sido cada vez mais empregados pelos agricultores (FURTADO et al., 2020). Essas novas ferramentas têm auxiliado por exemplo no controle de pragas, na identificação de deficiências nutricionais, na localização de falhas de plantio, no mapeamento da variabilidade, em processos de conservação do solo, na implementação de sistemas de irrigação mais eficientes e no desenvolvimento de novos protocolos laboratoriais (MOLIN et al., 2015). Essas "novas boas práticas" agrícolas só são possíveis graças a ampla gama de informações hoje compartilhada e compreendida, que fornecem assim subsídio para o manejo agrícola mais preciso (FERRAZ BARROS; MILAN, 2010). Neste contexto, dois conjuntos de ferramentas têm se mostrado poderosos aliadas na produção agrícola: as assinaturas espectrais e a inteligência artificial.

As assinaturas espectrais são eficientes na detecção de anomalias nas lavouras, uma vez que, plantas sadias exibem uma curva singular de energia no espectro eletromagnético (LIU 2006). DI LORENZO et al (2021), ao estudar assinaturas espectrais de vinhedos tanto por equipamento de contato quanto por imagem orbitais conseguiram distinguir diferentes cultivares para uva de mesa, além disso, conseguiram distinguir diferentes práticas de gestão entre plantações de mesma cultivar. VILLACRÉS et al. (2021) desenvolveram um índice de

vegetação através das assinaturas espectrais para analisar conteúdo de água na folha de eucaliptos e pinus, os resultados sugerem que, a partir do índice criado para água na folha outros índices espectrais podem ser recuperados com um rmse de até 0,02 e um r^2 de 0,77. E NETO et al. (2019), ao analisar assinaturas espectrais de cafezal ao longo dos anos conseguiu distinguir entre os anos qual foi o ano com maior produtividade no local, além disso, descreveu possíveis mecanismos fenológicos predominantes em cada ano que resultaram nas produtividades através da assinatura espectral.

Já a inteligência artificial (IA) é um conjunto de ferramentas que juntas são capazes de permitir que sistemas e computadores possam ver, aprender, ententer e decidir sobre situações as que estão expostas (ABDAR et al., 2021). YANG et al. (2021), ao estudarem formas de predizer a germinação de sementes de beterraba através do uso de imagens hiperespectrais conseguiram elaborar um modelo de predição onde foi obtida uma acurácia na predição de de 95,5% utilizando a curva espectral completa e utilizando apenas alguns cumprimentos de onda foi de 92,32%. E DE MEDEIROS et al. (2020), ao estudar sementes de gramínea forrageira desenvolveram dois modelos de classificadores de qualidade para sementes utilizando espectro do infravermelho próximo (NIR) (modelo 1) e imagens de raio-x (modelo 2), os resultados apresentaram germinação uma precisão de 82% (modelo 1) e 90% (modelo 2), já para vigor apresentaram uma precisão de 61% (modelo 1) e 68% (modelo 2), combinando os modelos a precisão foi de 85% para germinação e 62% para vigor.

Diante do exposto, objetivou-se atráves do presente estudo o desenvolvimento de duas ferramentas para auxiliar o setor agrícola brasileiro, a 1ª ferramenta “RESAP” com capacidade de antecipar a qualidade das lavouras através da análise da região espectral do *Red-Edge*, e a 2ª ferramenta “ANNSEED” com capacidade de antecipar a qualidade dos lotes de sementes através da análise de imagens multiespectrais.

REFERÊNCIAS

ABDAR, M. et al. A review of uncertainty quantification in deep learning: Techniques, applications and challenges. **Information Fusion**, v. 76, p. 243–297, 2021.

APROSOJA. Acompanhamento da safra de soja. **APROSOJA**, v. 12, p. 1–5, 2019. Disponível em: <http://www.conab.gov.br/OlalaCMS/uploads/arquivos/17_06_08_09_02_48_boletim_graos_junho_2017.pdf>.

COELHO, F.; LOVATTO, M.; SILVA, G.B.P. Soja: estresse caro. **Revista Cultivar**, n. 233, p. 22-26, out. 2018.

CONAB - COMPANHIA NACIONAL DE ABASTECIMENTO. Acompanhamento da Safra Brasileira de Grãos: Safra 2020/21. **BRASIL**, v. 8, p. 1–115, 2021.

DE MEDEIROS, A. D. et al. Machine learning for seed quality classification: An advanced approach using merger data from FT-NIR spectroscopy and x-ray imaging. **Sensors (Switzerland)**, v. 20, n. 15, p. 1–12, 2020.

DI LORENZO, R. et al. The use of vineyard spectral signatures to identify table grape cultivars. **Acta Horticulturae**. 1314, 197-204, 2021. Disponível em:<https://doi.org/10.17660/ActaHortic.2021.1314.25 >.

FARIAS, J. R. B.; NEPOMUCENO, A. L.; NEUMAIER, N. Circular Técnica 48: Ecofisiologia da soja. **EMBRAPA SOJA**, p. 9, 2007.

FEHR, W.R.; CAVINESS, C.E. **Stages of soybean development.** Ames: Agriculture and Home Economics Experiment Station Iowa State University of Science and Technology, 1977. (Special Report. 87)

FERRAZ BARROS, F.; MILAN, M. Qualidade operacional do plantio de cana-de-açúcar. **Bragantia**, v. 69, n. 1, p. 221–229, 2010.

FURTADO, G. D. F. et al. As geotecnologias e o agronegócio brasileiro: Um estudo de caso na produção de Soja. **Revista Brasileira de Agrotecnologia**, v. 10, n. 3, p. 88–96, 2020.

LIU, W.T.H. **Aplicações de Sensoriamento Remoto**. Campo Grande: Ed. UNIDERP, 908p., 2006.

MOLIN, J. P et al. **Agricultura de Precisão**. 1. ed. São Paulo: Oficina de Textos, v.1, 238p., 2015.

MORANDO, R. et al. Déficit Hídrico: Efeito Sobre a Cultura Da Soja. **Journal of Agronomic Sciences**, n. 3, p. 114–129, 2014.

MUNDSTOCK, C. M.; THOMAS, A. L. **Soja: fatores que afetam o crescimento e o rendimento de grãos**. 1. ed. Porto Alegre, RS: Departamento de Plantas de Lavoura da Universidade Federal do Rio Grande do Sul, 2005.

NETO, M. B. et al. UTILIZAÇÃO DE IMAGENS DO SENTINEL-2 PARA ANÁLISE DA VARIAÇÃO ESPECTRAL DA CULTURA DO CAFÉ NAS SAFRAS 2015/16 E 2016/17. (INPE, Ed.) In: XIX Simpósio Brasileiro de Sensoriamento Remoto, Santos, SP. **Anais**... Santos, SP: Galoá, 2019.

NEVES, M. F. et al. Estratégias para Conquistar a Posição de Fornecedor Mundial Sustentável de Alimentos, Bioenergia e outros Agro-Produto. **Revista Agronomia Brasileira**, v. 4, n. 2020, p. 1–6, 2020.

OECD/FAO. OECD-FAO **Agricultural Outlook 2020-2029**. Paris: OECD Publishing, 2020. Disponível em:<https://doi.org/10.1787/1112c23b-en.>. Acesso em: 16 de abr. 2021.

ONU - Organização das Nações Unidas. **World Population Prospects**, 2019. Disponível em: https://population.un.org/wpp/Publications/. Acesso em: 16 de abr. 2021.

SAATH, K. C. de O.; FACHINELLO, A. L. Crescimento da Demanda Mundial de Alimentos e Restrições do Fator Terra no Brasil. **Revista de Economia e Sociologia Rural**, v. 56, n. 2, p. 195–212, 2018.

VILLACRÉS, J. et al. Retrieval of vegetation indices related to leaf water content from a single index: A case study of eucalyptus globulus (labill.) and pinus radiata (d. don.). **Plants**, v. 10, n. 4, 2021.

YANG, J. et al. Hyperspectral prediction of sugarbeet seed germination based on gauss kernel SVM. **Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy**, v. 253, n. 74, p. 119585, 2021. Disponível em: <https://doi.org/10.1016/j.saa.2021.119585>.

2. RESAP: PROPOSTA DE MENSURAÇÃO DA INTENSIDADE DO "RED-EDGE POSITION" E SUA CORRELAÇÃO COM COMPONENTES DE PRODUÇÃO

RESUMO

A análise da região espectral do *Red-Edge* para culturas constitui-se como uma ferramenta ímpar no auxílio ao manejo de tratos culturais, bem como na antecipação do potencial produtivo. O presente estudo teve como objetivo a elaboração de ferramenta para mensurar a intensidade da "Posição do *Red-Edge*" através da abertura angular obtida pela reflectância e, determinar se tal ângulo possuia correlação significativa com as componentes de produção. Foram utilizadas assinaturas espectrais da folha central do 1º trifólio totalmente desenvolvido da soja, para isso foi utilizado o sensor hiperespectral FieldSpec® 3. Já as componentes de produção foram mensuradas em laboratório, segundo os protocolos estabelecidos pelo MAPA. A elaboração da ferramenta ocorreu através do desenvolvimento de uma nova rotina, onde foi combinado o método de extrapolação linear para determinar a posição do red-edge com relações trigonométricas no triângulo retângulo. Os valores obtidos em laboratório e através da ferramenta foram posteriormente submetidos a matriz de correlação. Os resultados mostraram que no estágio fenológico entre R3 e R4, os ângulos da Posição do *Red-Edge* mensurados foram fortemente correlacionadas as taxas de germinação dos lotes, e em início de R6 foram correlacionados ao número de vagens por planta e teor de umidade dos grãos. No geral, a ferramenta apresentou grande potencial para ser utilizada como uma alternativa a antecipação do potencial produtivo das lavouras.

Palavras-chave: Posição do Red-Edge; Infravermelho Próximo; Matriz de Correlação; Assinatura Espectral; Savitzky-Golay

ABSTRACT

The analysis of the Red-Edge zone for crops is a unique tool to help to assist in managing crop treatments, as well as, in anticipating the potential production. The present paper aimed to elaborate a tool to measure the "Red-Edge Position" intensity through the angular opening obtained by the reflectance and determine if this angle has a significant correlation with production components. It was used spectral signatures from the central leaf of the first fully developed soybeans' Trifolium, in addition, to obtain the signatures was used the FieldSpec® 3 hyperspectral sensor. The production components were measured in the laboratory, following the protocol established by MAPA. The tool elaboration occurred through the development of a new routine, where it was combined the linear extrapolation methodology to find the red-edge position with the right triangle trigonometric functions. The values obtained in the laboratory and through the tool were posteriorly submitted to a correlation matrix. The results demonstrated that in the R4 and R5 phenological stages, the measured Red-Edge position angles had a stronger correlation with soybean packs' germination rates, and at the beginning of the R6 stage were correlated to pods per plant and grain moisture content. Overall, the tool has shown a great potential to be used as an alternative in anticipating crop production potential.

Keywords: Red-Edge Position; Near infrared; Correlation matrix; Spectrum signature; Savitzky-Golay

2.1 INTRODUÇÃO

O uso de novas ferramentas para monitoramento e controle das lavouras pode auxiliar nas próximas décadas a promover a maior eficiência agrícola (FURTADO et al., 2020). Neste contexto, uma das ferramentas mais eficientes na detecção de anomalias nas lavouras é a assinatura espectral, uma vez que, plantas sadias exibem uma curva singular de energia no espectro eletromagnético (LIU 2006). As regiões mais importantes do espectro para avaliar a vegetação são: a região do visível (~400 a 700 nm), a região do infravermelho próximo (~780 a 2.500 nm) e a região do infravermelho médio (~2.500 a 5.000 nm) (MOLIN, 2015). Cada uma das regiões, está fortemente correlacionada a fatores dominantes que controlam a reflectância das folhas, na faixa do visível (RGB) os fatores são correlacionados aos pigmentos presentes, já na faixa do infravermelho próximo (NIR) os fatores são correlacionados a estrutura celular interna e, na faixa do infravermelho médio (SWIR) os fatores são correlacionados ao conteúdo presente de água na folha (THENKABAIL; LYON; HUETE, 2019). Ademais, existe uma interzona especial entre a região do visível e do infravermelho próximo, intitulada região do *Red-Edge,* ocorre que nas plantas essa região do espectro é considerada como sendo a que possui maior sensibilidade a mudanças ocorridas na vegetação (GITELSON et. al, 2011).

Segundo HORLER et al. (1983), a zona do red-edge é definida como a zona onde ocorre a mudança brusca na refletância da folha entre 680 a 750 nm, ou seja, entre a banda do vermelho e o infravermelho próximo. Essa zona possui características únicas, pois é resultado da interação óptica entre os mecanismos de absorção da clorofila e a composição do tecido da planta (HORLER; DOCKRAY; BARBER, 1983; SEAGER; FORD, 2002). Além disso, dentro dessa faixa existe um ponto que demarca a transição exata entre o vermelho e o infravermelho próximo, esse ponto é intitulado de REP (*Red-Edge Position*). Sendo definido como o ponto de maior declividade, ou seja, o ponto de inflexão entre as duas grandes regiões dentro da zona do *Red-Edge*, em síntese, tal ponto é caracterizado como o ponto mais sensível a mudanças sofridas pela planta (CHO; SKIDMORE, 2006; YONGJUN; JINGJING, 2016). Neste cenário, a detecção precoce de alterações fisiológica e estruturais através do REP, torna-se relevante e inovador para pesquisas sobre o funcionamento das plantas diante de estresses bióticos e abióticos (BELANCER; MILLER; BOYER, 1995; HUA et al., 2019; MCILWAINE; RIVAS CASADO; LEINSTER, 2019).

Ao longo dos anos, essa zona de transição do espectro eletromagnético, tem sido cada vez mais estudada pelos cientistas. Segundo análises do Scopus (2021), o termo "*Red-Edge*" foi registrado pela primeira vez em 1928, no total há 3.218 registros de artigos contendo o termo, porém, destes apenas 347 contém o termo "*Red-Edge Position*" sendo o primeiro registro em 1989. Ainda segundo as análises, a partir de 2003 houve um aumento exponencial das pesquisas sobre REP, especialmente em três áreas, "Terra e Ciências Planetárias" (17,8%), "Agricultura e Ciências Biológicas" (16,9%) e, "Engenharia" (12,8%) sendo as pesquisas na agricultura lideradas pela China seguida dos EUA e Alemanha (SCOPUS, 2021). A utilização da análise da zona do *Red-Edge* como ferramenta detecção de anomalias nas plantações, tem se mostrado uma poderosa aliada no monitoramento e controle das lavouras, exibindo resultados promissores na compreensão de como seria possível antecipar a qualidade das lavouras.

VOGELMANN et al. (1993) relataram detectar fortes correlações entre o teor de clorofila presente nas folhas de Carvalho Canadense e no REP, além disso, descrevem que diversos fatores abióticos possivelmente foram detectados nessa faixa do espectro ao decorrer do estudo, sendo necessário estudos mais aprofundados. MUTANGA e SKIDMORE (2007) buscando compreender a variabilidade da distribuição da fauna através do *Red-Edge* aplicaram diferentes doses de nitrogênio no Capim Buffei, os resultados indicaram que houve significativa alteração no teor de clorofila e que foi possível distinguir os níveis de nitrogênio através das alterações ocorridas no REP. Para EVRI et al. (2008) foi possível detectar em seus estudos para arroz fortes correlações entre o REP e os parâmetros mensurados, sendo positiva entre os teores de nitrogênio e clorofila, e negativa para o peso colhido de arroz. Segundo MARTINS e GALO (2015) foi possível diferenciar através do REP a cana-de-açúcar sadia das infectadas por nematoides e larvas do besouro migdolus [Migdolus fryanus], contudo, não foi possível diferenciar entre as infectadas. E BHOSLE e MUSANDE (2017) relataram ser possível através do REP diferenciar plantas doentes de saudáveis em diferentes culturas, sendo elas, amora, algodão e cana-de-açúcar.

Portanto, o objetivo deste presente capítulo foi elaborar uma metodologia de mensuração da intensidade do REP e determinar se os resultados obtidos possuiam correlação significativa com as componentes de produção.

2.2 MATERIAL E MÉTODOS

2.2.1 Área de estudo

A área de estudo está localizada no município de Piracicaba, no Estado de São Paulo, na latitude 22° 40,310’ S, longitude 47° 38,215’ W e altitude de 540 m (Fig. 1). O local apresenta solo Argissolo Vermelho (PV7), com presença de relevo ondulado (ROSSI, 2017).

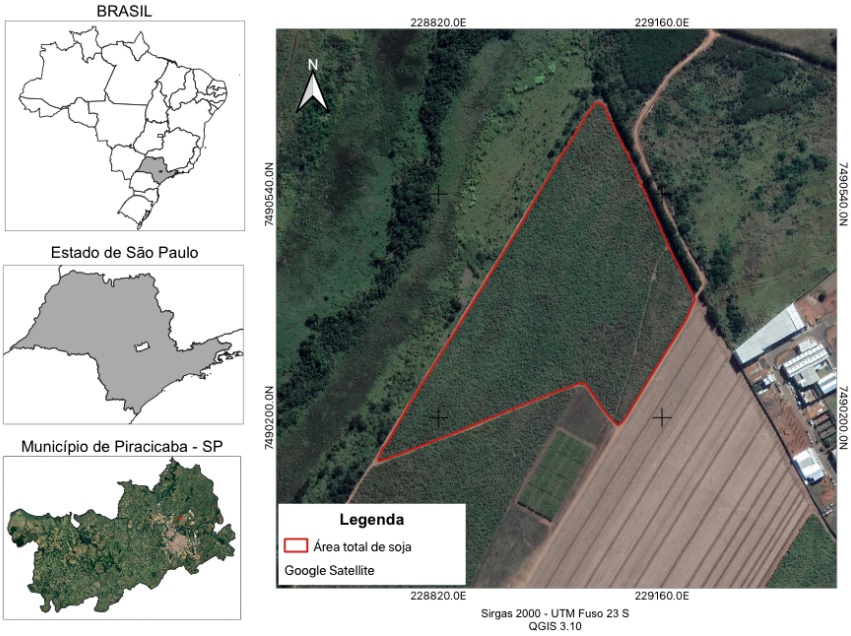


Figura 1: Mapa de localização

O clima da região foi classificado como mesotérmico subtropical (Cwa) segundo Köppen, com inverno (abril a setembro) apresentando clima frio, com baixa umidade relativa do ar e temperaturas próximas aos 18 ºC; já o verão (outubro a março) apresenta comportamento oposto e com temperaturas próximas aos 22 ºC; a precipitação média anual é de aproximadamente 1.250 mm (TAVARES, 2015).

2.2.2 Experimento

O experimento compreendeu uma área total de 1.225 m², inserida dentro de 11,6 ha cultivados com soja. Foram utilizadas duas cultivares, M6410 IPRO e M5947 IPRO, sendo essas subdivididas em dois tratamentos cada (Dose A e B), conforme a seguir tabela.

Tabela 1. Tratamentos empregados no local.

Cultivar	Dose	Tratamento
M6410 IPRO	A*	T1
	B**	T2
M5947 IPRO	A*	T3
	B**	T4

*Dose A: 4 doses do Biomax Turfa + 2 doses do Biomax 10 + 1 dose do Biomax Azum
**Dose B: 8,5 doses do Biomax Turfa + 5 doses do Biomax 10 + 2,5 dose do Biomax Azum

Para a implementação das doses optou-se pela técnica de coinoculação das sementes, durante o preparo da semeadora. Ocorre que a técnica escolhida apresentou melhores resultados em comparação a inoculação tradicional (MERCANTE et al., 2011; BATTISTI; SIMONETTI, 2015; GITTI, 2016; HUNGRIA; NOGUEIRA, 2017; BOSCHETTI; SIMONETTI, 2018; FILHO; SOUZA SILVA; SOUZA, 2018; MUNDIM et al., 2018; NOGUEIRA et al., 2018).

O plantio foi feito em sistema convencional, sendo o local de implementação selecionado a partir da análise de homogeneidade da condutividade elétrica aparente do solo (CEa) e da aplicação de calcário na área (calagem). Assim sendo, a semeadura ocorreu dia 13 de dezembro de 2018 nas primeiras horas da manhã, sendo utilizada uma população 14 pl.m^{-1} e espaçamento de 0,5 m entre linhas. Com emergência iniciada em 19 de dezembro de 2018.

2.2.3 Características agronômicas das cultivares

As cultivares selecionadas para o estudo são adaptadas ao Estado de São Paulo, sendo uma delas precoce e outra superprecoce, ambas com crescimento indeterminado e muito utilizadas pelo Brasil. Na tabela 2, encontram-se as características agronômicas de cada uma, sendo esses parâmetros utilizados posteriormente na compreensão dos resultados das análises.

Tabela 2. Características fenotípicas das cultivares utilizadas.

Características agronômicas	M6410 IPRO	M5947 IPRO
Ciclo	Precoce (127 dias)	Superprecoce (119 dias)
Grupo de maturidade	6,4	5,9
Arquitetura de planta	Semiereta	Ereta
Engalhamento	Bom	Moderado
Hábito de crescimento	Indeterminado	Indeterminado
Cor da flor	Roxa	Roxa
Cor da pubescência	Cinza	Cinza
Cor do hilo	Preto imperfeito	Preto imperfeito
Altura média da planta	86 cm	91 cm
Peso médio dos grãos	145g (1.000 grãos)	170g (1.000 grãos)
Acamamento	Resistente	Resistente
População	180 a 300 (mil/pl/ha)	240 a 300 (mil/pl/ha)

Fonte: Adaptado de Monsoy, 2019; Sementes Falcão, 2019.

A recomendação de semeadura ideal para ambas as cultivares em questão é entre 4ª semana de setembro à 4ª semana de outubro (MONSOY, 2019; SEMENTES FALCÃO, 2019).

2.2.4 Coleta de dados in situ e em laboratório

As cultivares em questão foram conduzidas estatisticamente separadas, assegurando dessa forma, que fosse minimizada a influencia genética nas análises.

2.2.4.1 In situ

a) Parâmetros climáticos

Os parâmetros climáticos diários de 2018 e 2019 utilizados, foram obtidos através do Posto Meteorológico localizado em Piracicaba-SP à uma distância de 3,5 km da área de estudo, sobre altitude de 546m. Sendo analisados os seguintes parâmetros: temperatura média (°C), insolação (horas), precipitação (mm) e evaporação (mm).

Os dados foram computados seguindo duas premissas: o ciclo planejado caso a semeadura ocorresse em 25 de setembro (conforme recomendação), e o ciclo ocorrido com semeadura em 13 de dezembro. Sendo estes acumulados a cada 6 dias, equivalente à um período, no total foram 19 períodos computados (ou seja, 114 dias de ciclo).

b) Fenologia e altura

Durante as coletas em campo foi identificado o estágio fenológico de cada uma das plantas amostradas, obedecendo a tabela 3 de classificação.

Tabela 3. Estágios da soja (crescimento determinado e indeterminado)

			Descrição
Estágio vegetativo		VE	Plântulas em emergência dos cotilédones.
		VC	Apresenta cotilédones totalmente abertos e expandidos.
		V1	Um par de folhas unifolioladas, ressalta-se que o nó cotiledonar não é considerado.
		V2	1º trifólio completamente desenvolvido, ou seja, dois nós.
		V3	Dois trifólios completamente desenvolvidos, ou seja, três nós.
		V4	Três trifólios completamente desenvolvidos, ou seja, quatro nós.
		V5	Quatro trifólios completamente desenvolvidos, ou seja, cinco nós.
		V6	Cinco trifólios completamente desenvolvidos, ou seja, seis nós.
		Vn	...
Estágio reprodutivo	**Estágio da flor**	R1	Uma flor aberta em qualquer nó na haste principal.
		R2	Uma flor aberta em um dos dois nós superiores na haste principal com folha completamente desenvolvida.
	Estágio da vagem	R3	Vagem de 0,5 a 2 cm em um dos quatro nós superiores na haste principal. Folhas V1 caminhando para a senescência (tom amarelado) e ainda apresenta floração.
		R4	Vagem completamente desenvolvida (>2 cm) em um dos quatro nós superiores na haste principal. Apresenta folhas V1 em completa senescência e ainda apresenta floração.
	Estágio do grão	R5	• Fase 3: Enchimento de grãos (26 a 50% de granação) em um dos quatro nós superiores na haste principal. Não apresenta mais flores. • Fase 5: Enchimento de grãos (76 a 100%) em um dos quatro nós superiores na haste principal.
		R6	Grão cheio ou completo em um dos quatro nós superiores na haste principal. Início da senescência (alguns trifólios amarelados).
	Estágio da maturação	R7	Início da maturação, folhas com forte presença de senescência (tom amarelada). Uma vagem com coloração madura na haste principal.
		R8	Maturação plena. Mais de 95% das vagens com coloração de madura, planta com perda total das folhas.

Fonte: adaptado de FEHR e CAVINESS (1977), FARIAS et al. (2007) e OLIVEIRA JUNIOR et al. (2016).

Da mesma forma, a altura de cada uma das plantas amostradas foi coletada, sendo a mensuração feita do ápice à base.

c) Espectros

Os espectros, ou assinaturas espectrais, foram adquiridos em cada uma das plantas amostradas a partir da análise da folha central do 1º trifólio totalmente desenvolvido do ápice à base, conforme recomendações (VELOSO et al., 2004; OLIVEIRA JUNIOR, 2016). No campo as folhas foram enumeradas conforme foram coletadas, garantindo assim que fossem analisadas na mesma ordem.

O equipamento utilizado nas análises foi o sensor hiperespectral FieldSpec® 3, fabricado pela empresa Analytical Spectral Devices Inc. (ASD, 2010). O sensor possui a capacidade de coletar dados de reflectância compreendidos entre 350 a 2.500 nm do espectro eletromagnético, sendo composto por duas faixas de resolução (3 nm e 10 nm) e dois intervalos de amostragem (1,4 nm e 2 nm). Para este presente estudo foram selecionados dados espectrais compreendidos entre 660 a 780 nm, e o sensor foi acoplado aos acessórios ASD

Plant Probe e Leaf Clip que permite o completo isolamento da área da folha, garantindo assim que a única fonte de luz incidente sobre a mesma seja a proveniente do sensor (ASD, 2010; CRUSIOL, 2017).

I. II.

Figura 2. I. FieldSpec® 3 Jr (ASD, 2010); II. Plant Probe com Leaf Clip.

2.2.4.2 Laboratório

As coletas dessa etapa foram conduzidas, no Laboratório de Análise de Sementes, do Departamento de Produção Vegetal – USP/ESALQ, segundo orientações do Ministério da Agricultura, Pecuária e Abastecimento (BRASIL, 2009). Neste contexto, visando a preservação das características de cada planta coletada, a colheita da área foi realizada de maneira manual evitando assim danos mecânicos nos indivíduos (sementes e plantas). No total foram 10 plantas coletadas aleatoriamente para cada tratamento, totalizando 40 plantas, sendo mensuradas as seguintes componentes de produção: número de vagens por planta, massa seca (g), teor de umidade (%), peso de mil grãos (g), taxa de germinação (%).

2.2.5 Obtenção da intensidade do REP

Após a coleta das curvas espectrais a partir das folhas, as mesmas foram plotadas de maneira individual para a verificação da qualidade com relação à presença de ruídos e feição padrão de dados de vegetação (LIU, 2006; MUSTAFA, 2015; THENKABAIL; LYON; HUETE, 2019).

A partir das observações foi aplicado o filtro de ruídos Savitzky-Golay em 1ª ordem com 5 pontos de suavização obedecendo estritamente o equacionamento exposto por MARK e WORKMAN JR. (2004). Na sequência foi feita a 1ª derivada em cada uma das curvas, seguindo rotinas aplicadas em outros estudos (ADAMS, 1995; EVRI; AKIYAMA; KAWAMURA, 2008; PANERO et al., 2018; MCILWAINE; RIVAS CASADO; LEINSTER, 2019). Para essa transformação de dados foi feito o equacionamento da 1ª derivada por Savitzky-Golay com 5 pontos obedecendo estritamente o equacionamento exposto por MARK e WORKMAN JR. (2004).

Posteriormente, foi aplicado o método de extrapolação linear proposto por CHO e SKIDMORE (2006) para determinar a posição exata do REP. O método consiste na extrapolação linear de linhas retas entre a saída do vermelho distante (680 a 700 nm) e do NIR (725 a 760 nm) do espectro em 1ª derivada, a intersecção entre as linhas demarca o REP. Esse método foi desenvolvio como uma alternativa de obtenção do REP de maneira menos complexa, contudo, eficaz sendo os resultados tão exatos quantos os obtidos atráves de técnica

tradicionais que utilizam equacionamentos complexos (interpolação linear, polinomial e ajuste gaussiano invertido).

Uma vez determinados os REPs, ou seja, o posicionamento do ponto de inflexão, os valores de reflectância relativos a cada um foram separados, sendo essa a reflectância a incidente no momento em que ocorreu a inflexão entre o vermelho e infravermelho próximo. Para critérios comparativos as reflectâncias foram dispostas em forma de triângulo retângulo (de mesma base e alturas diferentes) e seus ângulos de abertura mensurados através das regras trigonométricas.

Essa metodologia de comparação entre as aberturas angulares produzidas pela intensidade da reflectância recebeu pela presente autora deste trabalho o nome de **RESAP** (**R**ed-**E**dge's **S**tress **A**ngle **P**rotractor), sendo o seu principal objetivo facilitar a comparação e catalogar anomalias nas vegetações ao longo dos anos. Segue figura esquemática da metodologia RESAP.

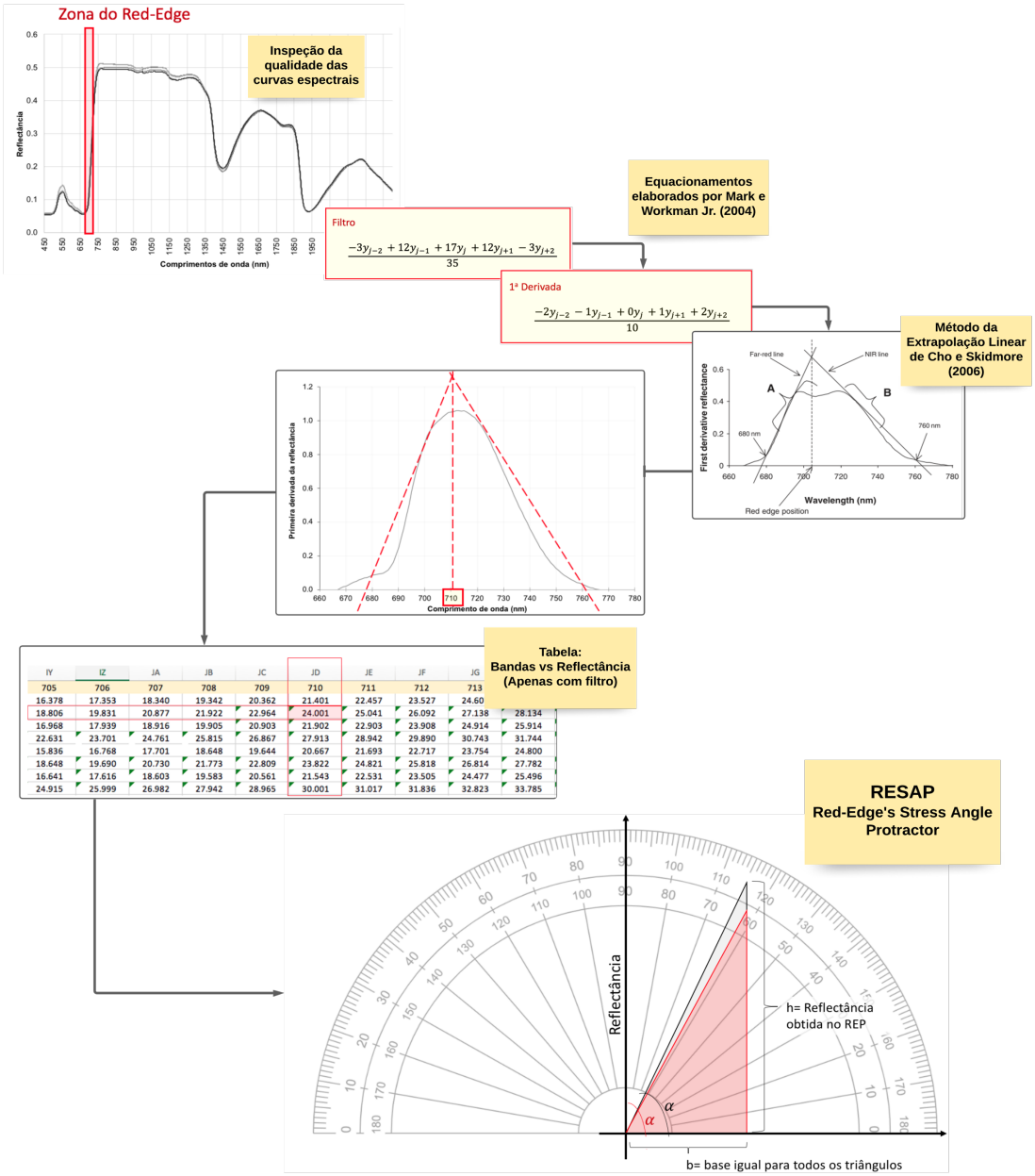


Figura 3. RESAP (Red-Edge's Stress Angle Protractor)

Os resultados obtidos em laboratório e RESAP, foram submetidos a análise de correlação através da aplicação da matriz de correlação para analisar o quanto correlacionado os dados eram e como eles estavam relacionados, em outras palavras, a intensidade e direção da relação entre eles. Em suma, a intensidade dessa relação pode ir de 0 a 1, sendo 1 uma correlação muito forte, e se ambas as variáveis aumentam ou diminuem recebem sinal positivo, caso contrário, ou seja, uma aumenta e outra diminui recebem sinal negativo.

2.2.6 Análise de dados

Os resultados obtidos, foram submetidos ao teste de normalidade de Anderson-Darling via simulação de Monte Carlo (RAZALI e WAH, 2011; YAP e SIM, 2011), tendo como intuito avaliar se os dados possuíam uma distribuição normal. A partir do teste de normalidade, foi efetuada a análise de variância das cultivares seguido pelo teste par-a-par, sendo analisado o tratamento no decorrer dos dias e os tratamentos entre si para o mesmo dia. Ou seja, uma análise de variância conjunta pelos dias e pelos tratamentos para cada cultivar.

Por último, foi realizada a matriz de correlação com os resultados das componentes de produção e das curvas espectrais, sendo os valores de correlação exibidos no triângulo inferior e as probabilidades no triângulo superior.

Todas as etapas compreendidas na metodologia seguem ilustradas abaixo.

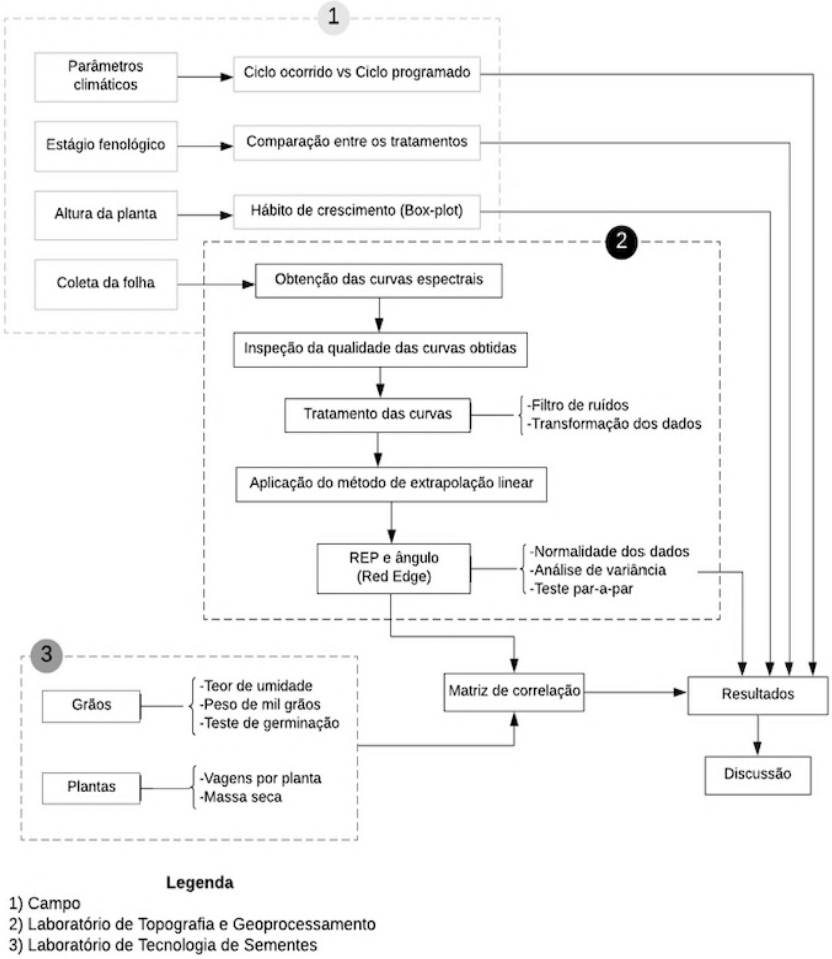


Figura 4. Etapas empregadas.

2.3 RESULTADOS E DISCUSSÃO

A reflectância resultante obtida a partir das assinaturas espectrais ilustra uma diferenciação entre dias e entre tratamentos. Ocorre que as cultivares exibiram comportamento inverso de acordo com a dose de inoculante recebida, posteriormente sendo evidente nas análises das componentes de produção que uma delas obteve melhor desempenho dentre todas.

Ao analisar a plotagem das assinaturas espectrais resultantes, foi possível perceber que as curvas dos três primeiros dias de coleta (09, 14 e 20 de fevereiro) assemelham-se dentro de cada tratamento (T1, T2, T3, T4), e a entrada na zona do *Red-Edge* (entorno de 680 a 690 nm) ocorre de maneira análoga havendo sobreposição das curvas em diversas bandas do espectro eletromagnético. Já a curva do quarto dia (23 de março) difere-se das demais coletas dentro do mesmo tratamento, sendo a mesma afastada das outras. Ocorre que na última coleta as plantas estavam em estágio fenológico avançado (R6-R7, conforme exposto nos Apêndice B e C). Nesses estágios o processo de degradação da clorofila se inicia, portanto, as plantas estavam caminhando para a senescência e aproximando-se do ponto de colheita (FARIAS; NEPOMUCENO; NEUMAIER, 2007). Dessa forma, as curvas espectrais do dia 23 de março em todos os tratamentos exibiram o fenômeno intitulado de "*blue shift of the red-edge*". Este ocorre quando há uma redução ou ausência de clorofila na planta, o que ocasiona uma diminuição ou parada na absorção de energia pela mesma, causando deslocamento na assinatura espectral, sendo na região do visível para cima e na região do red-edge para a esquerda (HORLER; DOCKRAY; BARBER, 1983; CARTER, 1993; MUTANGA; SKIDMORE, 2007; EVRI; AKIYAMA; KAWAMURA, 2008). Por consequência dessa redução ou ausência de clorofila, a planta passa a apresentar tons amarelados em suas folhas (TAIZ & ZEIGER, 2004; MODESTO; SIQUEIRA, 1981). Portanto, este dia apresentou uma assinatura espectral resultante diferente devido primordialmente ao estágio fenológico. Segue a plotagem das assinaturas espectrais resultantes, por tratamento (Figura 5).

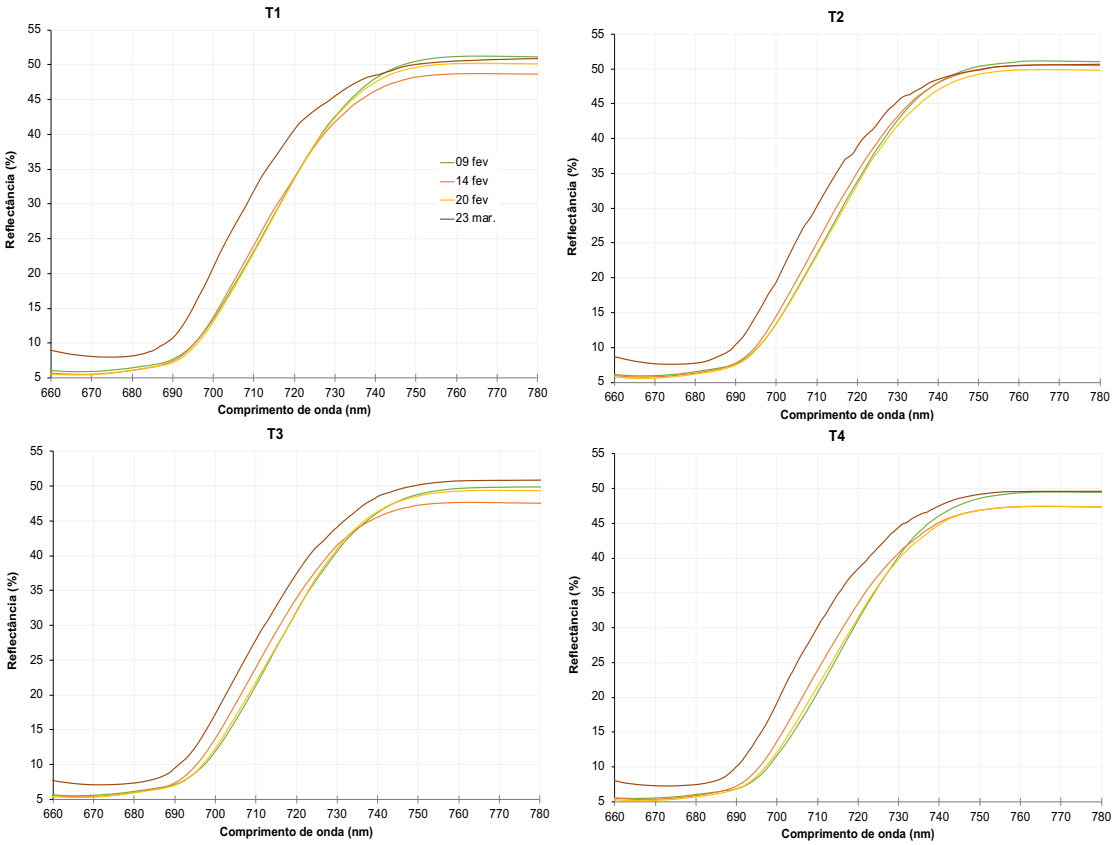


Figura 5. Assinaturas espectrais resultantes da soja.

A fim de particularizar ainda mais as assinaturas espectrais resultantes da primeira a terceira coleta, foram calculadas as derivações dos espectros em 1ª ordem. Essa prática facilita a visualização e manuseio dos dados, pois espectros derivados enfatizam melhor pequenas diferenças, ademais, ressaltam mudanças na inclinação que dificilmente seriam detectadas em ordem zero (ADAMS, 1995). O resultado dessa derivação mostrou que o dia 23 de março (última coleta) ainda permaneceu muito distinto das demais, porém essa prática permitiu diferenciar mais um dia dentro das coletas. O dia 14 de fevereiro (segunda coleta), distinguiu-se dos outros dois dias de coleta em todos os tratamentos. Já os dias 09 e 20 de fevereiro ilustraram uma tendência a permanecerem sobrepostos mesmo depois da derivação, sendo diferenciáveis apenas em T4.

Sobre o dia 14 de fevereiro, o que se observou nos dados complementares (em Apêndice A, B e C), foi uma transição entre estágios somada à uma inversão climática após a primeira coleta, permanecendo dessa forma até a terceira coleta. Processo este em que a temperatura tornou-se ideal (entre 20 a 30 °C), além disso houve uma queda brusca na insolação acumulada por período, e teve início um breve período de chuvas. Por serem cultivares com hábito de crescimento indeterminado, pode-se inferir que durante o período entorno do dia 14 de fevereiro as plantas estavam vegetando. O que resultou em um intenso crescimento em cada um dos tratamentos, média de 11 cm em T1, 6 cm em T2, 5 cm em T3 e 5,5 cm em T4 (conforme exposto no Apêndice E). De acordo com FEHR e CAVINESS (1977), TAIZ & ZEIGER (2004) e MODESTO e SIQUEIRA (1981), plantas que apresentam o hábito de crescimento indeterminado demonstram uma tendência a diminuir o crescimento em determinados estágios uma vez que toda sua a energia é translocada para a produção de descendentes. Processo

esse que pode ser intensificado em função de estresse, como por exemplo, estresse abiótico podendo ser decorrente das adversidades climáticas[1] (MUNDSTOCK; THOMAS, 2005; FARIAS; NEPOMUCENO; NEUMAIER, 2007), conforme ocorrido durante a experimentação em questão. Já nas coletas do dia 09 e 20 de fevereiro as plantas estavam mudando de estágio fenológico, no primeiro momento finalizando a formação das vagens e iniciando o processo de enchimento de grãos, e já no terceiro finalizando o enchimento de grãos e iniciando a degradação da clorofila, conforme em Apêndice B e C. Ademais, é essencial mencionar que o T4 dentre todos foi o que exibiu uma distinta separação entre as curvas espectrais resultantes das coletas. Portanto, considerando que todas as plantas estavam sobre mesmas condições climáticas e os tratamentos aleatorizados no campo, a única diferença existente foram as doses de inoculante fornecidas aos tratamentos, assim sendo, pode-se inferir após observar as componentes de produção em Apêndice D que T4 estaria absorvendo de maneira mais eficiente os nutrientes disponíveis, contudo análises específicas entre a interação do genótipo e os mecanismos de absorção de macronutrientes são necessárias.

Segue abaixo a 1ª derivada das assinaturas espectrais resultantes por tratamento, em suma, infere-se: no dia 23 de março o processo de senescência, no dia 14 de fevereiro o crescimento, no dia 09 e 20 de fevereiro a finalização dos devidos estágios fenológicos, e o comportamento particularizado de T4.

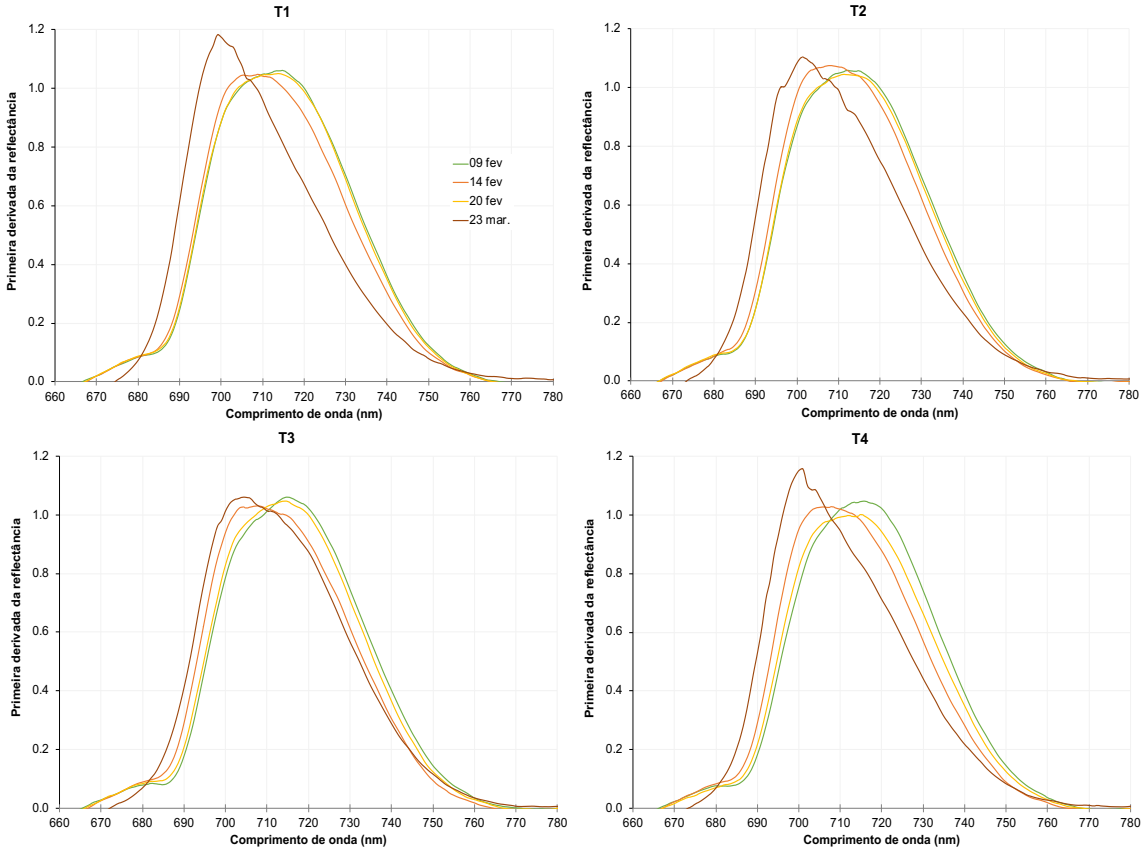


Figura 6. 1ª Derivada das assinaturas espectrais resultantes.

Nota: [1] As adversidades supracitadas foram provenientes de altas temperaturas médias, dias longos e noites curtas, alta taxa de insolação e, baixo volume de precipitações no local. Conforme exposto no Apêndice A deste presente capítulo.

Após as análises visuais das curvas espectrais (Fig. 5 e Fig. 6), foram calculados os REP's, identificadas as reflectâncias correspondentes aos REP's e os ângulos de abertura de cada uma das curvas. Os resultados foram submetidos ao teste de normalidade de Anderson-Darling via simulação de Monte Carlo com 10.000 permutações, os dados apresentaram uma distribuição anormal. Assim sendo, optou-se pelo uso de testes estatísticos não-paramétricos, analisando o tratamento pelos dias e comparando os tratamentos no dia (T1 vs T2 e T3 vs T4). Na análise do tratamento, foi utilizado o teste de Kruskall-Wallis (valor p menor ou igual a 0,05) para avaliar se existiam diferenças estatísticas significativas entre as medianas, seguido pelo teste par-a-par de Mann-Whitney sinalizando as medianas consideradas estatisticamente diferentes. E para a comparação entre tratamentos, foi utilizado o teste U de Mann-Whitney (*two-sample tests)* para avaliar se existia diferença estatística significativa entre as medianas dos tratamentos no dia em questão, sendo avaliado separadamente cada um dos dias de coleta.

Os resultados apresentados nas tabelas a seguir, mostram que apesar das cultivares serem distintas e existirem doses de inoculante diferentes, elas tendem a apresentar um comportamento semelhante, o qual permite inferir que este seja o comportamento característico da soja. Ademais, infere-se que de maneira geral este comportamento está vinculado predominantemente ao estágio fenológico apresentado pelas plantas nos dias das coletas (HORLER; DOCKRAY; BARBER, 1983). Os resultados confirmaram a discussão realizada anteriormente neste presente estudo ao analisar apenas os gráficos de assinatura espectral e primeira derivada da assinatura. Ou seja, eles confirmam a existência de três grupos de reflectância, sendo eles as assinaturas do dia 09 e 20 de fevereiro, as do dia 14 de fevereiro e as do dia 23 de março. Se deduz que as pequenas diferenças numéricas apresentadas nos valores de reflectância e ângulo, podem ser indícios de absorção mais eficiente dos nutrientes disponíveis, fato fortificado posteriormente pela matriz de correlação entre as componentes de produção e os ângulos de abertura de cada um dos tratamentos. Segue abaixo os resultados das análises estatísticas empregadas no ângulo de abertura obtido por meio do RESAP.

Tabela 4. Resultados estatísticos do ângulo alfa do RESAP para T1 e T2.

	Tratamento	**REP (banda)**	**Reflectância (%)**	**RESAP α (º)**	**Comparação entre T1 e T2 (p)**
09 fev	**T1**	710	23,229	$66,71^{a}$	0,0866
	T2	710	23,321	$66,79^{A}$	
14 fev	**T1**	706	19,853	$63,27^{b}$	0,2394
	T2	706	20,637	$64,15^{B}$	
20 fev	**T1**	710	22,988	$66,49^{a}$	0,3452
	T2	710	23,113	$66,60^{B}$	
23 mar	**T1**	698	18,376	$61,45^{b}$	0,5449
	T2	698	17,604	$60,40^{C}$	

Medianas seguidas da mesma letra não diferem entre si pelo teste de Kruskall-Wallis seguido da comparação par-a-par de Mann-Whitney a $p \leq 0,05$.

Tabela 5. Resultados estatísticos do ângulo alfa do RESAP para T3 e T4.

	Tratamento	REP (banda)	Reflectância (%)	RESAP α (º)	Comparação entre T3 e T4 (p)
09 fev	**T3**	712	23,527	66,97ª	0,8126
	T4	712	22,717	66,24^{A}	
14 fev	**T3**	706	19,831	63,24^{b}	0,0970
	T4	706	19,690	63,08^{C}	
20 fev	**T3**	710	21,902	65,46ª	0,0004*
	T4	710	21,543	65,10^{B}	
23 mar	**T3**	702	19,410	62,74^{b}	0,0040*
	T4	698	16,788	59,22^{C}	

Medianas seguidas da mesma letra não diferem entre si pelo teste de Kruskall-Wallis seguido da comparação par-a-par de Mann-Whitney a $p \leq 0,05$.
* Houve diferença significativa entre os tratamentos no dia analisado.

Como última análise, os ângulos resultantes de cada um dos tratamentos foram confrontados com os resultados obtidos de componentes de produção, através de matriz de correlação pelo coeficiente de Spearman, sendo este utilizado para análise de dados não-paramétricos. Dessa forma, pode-se analisar o quanto correlacionado os dados estão e como eles estão relacionados, em outras palavras, a intensidade e direção da relação dos dados. Em suma, a intensidade dessa relação pode ir de 0 a 1, sendo 1 uma correlação muito forte, e se ambas as variáveis aumentam ou diminuem recebem sinal positivo, caso contrário, ou seja, uma aumenta e outra diminui recebem sinal negativo.

Neste contexto, foi detectado que a taxa de germinação foi a componente de produção que apresentou maior correlação com os ângulos do *Red-Edge* em três dias de coleta, sendo no dia 09 de fevereiro uma correlação positiva muito forte, no dia 14 de fevereiro uma correlação positiva fraca, e no dia 23 de março uma correlação positiva forte. Dessa forma, quanto maior o ângulo encontrado maior foi a possibilidade de germinação do lote colhido, este padrão permaneceu quando as cultivares foram analisadas separadamente. Já o dia 20 de fevereiro, exibiu uma correlação negativa muito forte com o número de vagens e teor de umidade. Ou seja, quanto maior o ângulo do red-edge menor o número de vagens, ou teor de umidade dos grãos colhidos.

Tabela 6. Matriz de correlação segundo coeficiente de Spearman, entre componentes de rendimento e RESAP.

	Nº vagens por planta	Massa seca (g)	Peso de mil grãos (g)	Taxa de germinação (%)	Teor de umidade (%)	RESAP alfa (°) – 09 fev	RESAP alfa (°) – 14 fev	RESAP alfa (°) – 20 fev	RESAP alfa (°) – 23 mar
Nº vagens por planta		-	-	-	-	0,33	0,67	**0,16**	0,42
Massa seca (g)	-		-	-	-	0,56	0,89	0,24	0,55
Peso de mil grãos (g)	-	-		-	-	0,84	0,77	0,53	0,98
Taxa de germinação (%)	-	-	-		-	**0,05**	0,53	0,70	0,23
Teor de umidade (%)	-	-	-	-		0,62	0,78	**0,16**	0,67
RESAP alfa (°) – 09 fev	-0,67	-0,44	-0,16	**0,95**	-0,38		-	-	-
RESAP alfa (°) – 14 fev	-0,33	-0,11	-0,23	0,47	-0,22	-		-	-
RESAP alfa (°) – 20 fev	**-0,84**	-0,76	0,47	0,30	**-0,84**	-	-		-
RESAP alfa (°) – 23 mar	-0,59	-0,45	-0,02	0,78	-0,34	-	-	-	

Observação: Valores de correlação estão dispostos no triângulo inferior, e no triângulo superior encontram-se as probabilidades de duas caudas (two-tailed) de não haver correlação entre as variáveis (ou seja, quanto mais próximo de zero maior será a probabilidade de correlação entre os dados).

Classificação dos valores de correlação:
(+/-) 0,90 a 1 – Muito Forte
(+/-) 0,7 a 0,89 – Forte
(+/-) 0,50 a 0,69 – Moderada
(+/-) 0,01 a 0,49 – Fraca

Classificação dos valores de probabilidade (two-tailed):
(+) 0,01 a 0,49 – Muito Alta
(+) 0,50 a 0,69 – Alta
(+) 0,7 a 0,89 –Moderada
(+) 0,90 a 1 – Baixa

2.4 CONCLUSÕES

Os ângulos obtidos atráves da metodologia RESAP mostraram correlação significativa positiva muito forte com a taxa de germinação.

O monitoramento do fenômeno intitulado de *"blue shift of the red-edge"* pode auxiliar no planejamento da colheita de áreas de soja.

Espectros de folhas no estágio fenológico entre R3 e R4 para a soja, mostrou ter a capacidade de antecipar a taxa de germinação.

Agradecimentos

O presente trabalho foi realizado em parte com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Contando também com apoio da FAPESP pela aquisição do equipamento FieldSpec® 3 Jr através do projeto de pesquisa nº 2013/22435-9.

REFERÊNCIAS

ADAMS, M.J. **Chemometrics in analytical spectroscopy.** Wolverhampton: Bookcraft, 1995.

ANALYTICAL SPECTRAL DEVICES, INC. (ASD INC.). **FieldSpec 3 User Manual.** Boulder: ASD Inc., 2010.

BATTISTI, A. M.; SIMONETTI, A. P. M. Inoculação e coinoculação com Bradyrhizobium japonicum e Azospirillum brasilense na cultura da soja. **Revista Cultivando o Saber**, v. 8, p. 294–301, 2015.

BELANCER, M. J.; MILLER, J. R.; BOYER, M. G. Comparative relationships between some red edge parameters and seasonal leaf chlorophyll concentrations. **Canadian Journal of Remote Sensing**, v. 21, n. 1, p. 16–21, 1995.

BHOSLE, K.; MUSANDE, V. Red edge point detection for mulberry leaf. In: Proceedings - 1st International Conference on Intelligent Systems and Information Management, ICISIM 2017, Aurangabad, India. **Anais**... Aurangabad, India: IEEE Xplore, 2017.

BOSCHETTI, E. L.; SIMONETTI, A. P. M. Influência da coinoculação de Bradyrhizobium e Azospirillum no desenvolvimento inicial da soja Influence of coadjulation of bradyrhizobium and azospirillum on early soybean development Introdução A produção de soja (Glycine max) é uma das atividades agr. **Revista Cultivando o Saber**, v. Edição Esp, p. 44–52, 2018.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. **Regras para análise de sementes.** Brasília: MAPA, 2009.

Carter, Gregory A. Primary and Secondary effects of the water content of the spectral reflectance of leaves. **American Journal of Botany**, v.74, n.7, p.916-924, 1993.

CHO, M. A.; SKIDMORE, A. K. A new technique for extracting the red edge position from hyperspectral data: The linear extrapolation method. **Remote Sensing of Environment**, v. 101, n. 2, p. 181–193, 2006.

CRUSIOL, L. G. T. **Dados multi e hiperespectrais da cultura da soja (Glycine Max L .) e sua relação com doses de gesso e calcário no solo**. 2017. Universidade Estadual de Maringá, 2017.

EVRI, M.; AKIYAMA, T.; KAWAMURA, K. Spectrum analysis of hyperspectral red edge position to predict rice biophysical parameters and grain weight. **Journal of the Japan society of photogrammetry and remote sensing**, v. 47, n. 2, p. 4–15, 2008.

FARIAS, J. R. B.; NEPOMUCENO, A. L.; NEUMAIER, N. Circular Técnica 48: Ecofisiologia da soja. **EMBRAPA SOJA**, p. 9, 2007.

FEHR, W.R.; CAVINESS, C.E. **Stages of soybean development.** Ames: Agriculture and Home Economics Experiment Station Iowa State University of Science and Technology, 1977. (Special Report. 87)

FILHO, J. M.; SOUZA SILVA, C. H.; SOUZA, J. E. B. DESEMPENHO AGRONÔMICO E PRODUTIVIDADE DA CULTURA DA SOJA COM A CO- INOCULAÇÃO DE Bradyrhizobium E Azospirillum brasilense. **Ipê Agronomic Journal**, v. 2, p. 48–59, 2018.

FURTADO, G. D. F. et al. As geotecnologias e o agronegócio brasileiro: Um estudo de caso na produção de Soja. **Revista Brasileira de Agrotecnologia**, v. 10, n. 3, p. 88–96, 2020.

GITELSON, A.A.; THENKABAIL, P.S.; LYON, J.G.; HUETE, A. **Remote sensing estimation of crop biophysical characteristics at various scales**. Hyperspectral remote sensing of vegetation. CRC Press, p. 329–358, New York, 2011.

GITTI, D. D. C. Inoculação e coinoculação na cultura da soja. **Fundação MS**, p. 15–24, 2016. Disponível em: <http://www.fundacaoms.org.br/base/www/fundacaoms.org.br/media/attachments/219/219/58347f37dea31067d8493f1a2d6afa7d8773fcd4970e0_01-inoculac-a-o-e-coinoculac-a-o.pdf>.

HORLER, D. N. H.; DOCKRAY, M.; BARBER, J. The red edge of plant leaf reflectance. **International Journal of Remote Sensing**, v. 4, p. 273–288, 1983.

HUA, Q. et al. Leaf spectral responses of Poa crymophila to nitrogen deposition and climate change on Qinghai-Tibetan Plateau. **Agriculture, Ecosystems and Environment**, v. 284, n. May, 2019.

HUNGRIA, M.; NOGUEIRA, M. A. COINOCULAÇÃO DA SOJA COM Bradyrhizobium E Azospirillum : UMA TECNOLOGIA AMBIENTALMENTE SUSTENTÁVEL. **XXXVI Reunião de Pesquisa de Soja**, p. 203–205, 2017.

LIU, W.T.H. **Aplicações de Sensoriamento Remoto**. Campo Grande: Ed. UNIDERP, 908p., 2006.

MARK, H.; WORKMAN JUNIOR, J. Derivatives in Spectroscopy: Part IV - Calibrating with Derivatives. **Spectroscopy**, v. 19, n. 1, p. 44–51, 2004.

MARTINS, G. D.; TRINDADE GALO, M. D. L. B. CARACTERIZAÇÃO ESPECTRAL DA CANA-DE-AÇÚCAR INFECTADA POR NEMATOIDES E MIGDOLUS FRYANUS POR ESPECTRORRADIOMETRIA DE CAMPO. **Boletim de Ciencias Geodesicas**, v. 21, n. 4, p. 783–796, 2015.

MCILWAINE, B.; RIVAS CASADO, M.; LEINSTER, P. Using 1st Derivative Reflectance Signatures within a Remote Sensing Framework to Identify Macroalgae in Marine Environments. **Remote Sensing**, v. 11, n. 6, p. 704, 2019.

MERCANTE, F. M. et al. Comunicado Técnico 169: Estratégias para Aumentar a Eficiência de Inoculantes Microbianos na Cultura da Soja. **Embrapa Agropecuária Oeste**, p. 4, 2011.

MODESTO, Z.M.; SIQUEIRA, N.J.B. **Currículo de estudos de biologia:** botânica. São Paulo: EPU, 1981.

MOLIN, J. P. et al. **Agricultura de Precisão**. 1. ed. São Paulo: Oficina de Textos, v.1, 238p., 2015.

MONSOY. Variedades Monsoy. 2019. Disponível em: <http://www.monsoy.com.br/variedades> Acesso em: 20 de maio de 2019.

MUNDIM, L. et al. Coinoculação de Azospirillum brasilense e Bradyrhizobium via sementes de soja no cerrado. **Global Science and Technology**, v. 11, n. 3, p. 10–19, 2018.

MUNDSTOCK, C. M.; THOMAS, A. L. **Soja: fatores que afetam o crescimento e o rendimento de grãos**. 1. ed. Porto Alegre, RS: Departamento de Plantas de Lavoura da Universidade Federal do Rio Grande do Sul, 2005.

MUSTAFA, N. et al. Quantification of oil palm tree leaf pigment (chlorophyll a) concentration based on theis age. **Jurnal Teknologi (Sciences & Engineering)**, 75:11, p. 129-134, 2015.

MUTANGA, O.; SKIDMORE, A. K. Red edge shift and biochemical content in grass canopies. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 62, n. 1, p. 34–42, 2007.

NOGUEIRA, M. A. et al. Circular técnica 143: Ações de transferência de tecnologia em inoculação/coinoculação com Bradyrhizobium e Azospirillum na cultura da soja na safra 2017/18 no estado do Paraná. **EMBRAPA SOJA**, p. 1–6, 2018.

OLIVEIRA JUNIOR, A. de; CASTRO, C. de; PEREIRA, L. R.; DOMINGOS, C. da S. **Estádios fenológicos e marcha de absorção de nutrientes da soja.** Londrina: Embrapa Soja, 2016.

PANERO, J. S. et al. Separation of Cultivars of Soybeans by Chemometric Methods Using Near Infrared Spectroscopy. **Journal of Agricultural Science**, v. 10, n. 4, p. 351–360, 2018.

RAZALI, N. M.; WAH, Y. B. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. **Journal of Statistical Modeling and Analytics**, Malaysia, v. 2, n. 1, p. 21-33, 2011.

ROSSI, M. **Mapa Pedológico do Estado de São Paulo: revisado e ampliado**. 1. ed. São Paulo, SP: Instituto Florestal, 2017.

SEAGER, S.; FORD, E. B. The Vegetation Red Edge Spectroscopic Feature as a Surface Biomarker. In: Astrophysics of Life: Proceedings of the Space Telescope Science Institute Symposium, Baltimore, Maryland. **Anais**... Baltimore, Maryland: Cambridge University Press, 2002. Disponível em: <http://arxiv.org/abs/astro-ph/0212550>.

SEMENTES FALCÃO. Guia de sementes de soja 2016. Disponível em: <http://www.sementesfalcao.agr.br/xdown/GuiaSoja2016SF.pdf> Acesso em: 20 de maio de 2019.

TAIZ, L. & ZEIGER, E. 2004. **Fisiologia Vegetal.** 3ª ed. Porto Alegre: Artmed. 719p.

TAVARES, T.R.; FIORIO, P.R.; MARTELLO, M.; STRABELI, T.F.; BARROS, P.P.S. Observação do Normalized Difference Vegetation Index (NDVI) de diferentes tipos de vegetação nativa no período de estiagem de 2011 e 2014, em Piracicaba-SP. **Revista Hipótese**, v.1, n.2, p.77-91, 2015.

THENKABAIL, P.S.; LYON, J.G.; HUETE, A. **Hyperspectral Remote Sensing of Vegetation.** CRC PRESS. New York, 2019.

VELOSO, C.A.C. et al. **Amostragem de plantas para análise química.** Comunicado Técnico. Belém: Embrapa Amazônia Oriental, 2004.

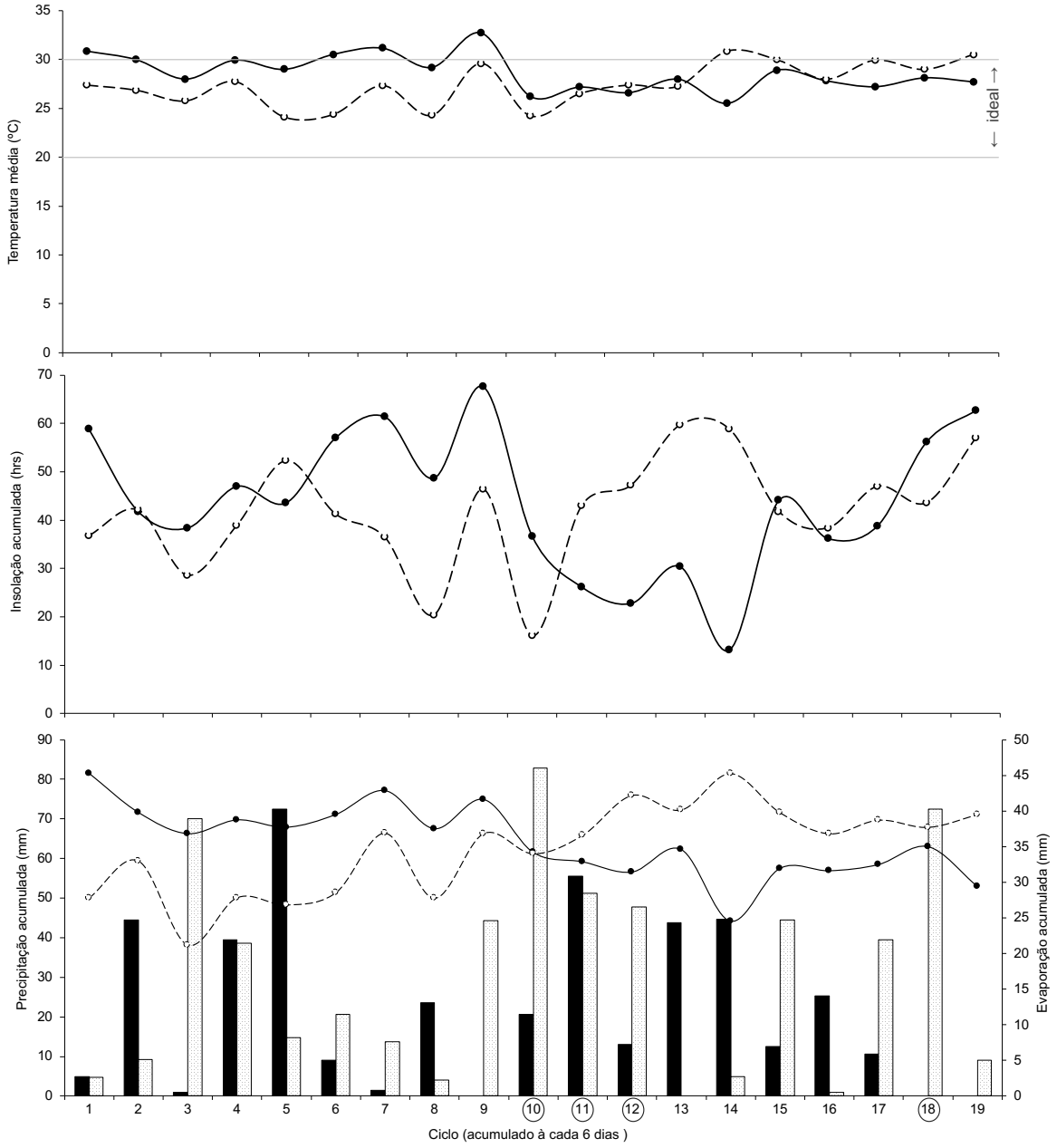
VOGELMANN, J. E.; ROCK, B. N.; MOSS, D. M. Red edge spectral measurements from sugar maple leaves. **International Journal of Remote Sensing**, v. 14, n. 8, p. 1563–1575, 1993.

YAP, B. W.; SIM, C. H. Comparisons of various types of normality tests. **Journal of Statistical Computation and Simulation**, London, v. 81, n. 12, p. 2141-2155, 2011.

YONGJUN, D.; JINGJING, Z. Extraction of REPs from leaves reflectance spectrum for estimation of chlorophyll content. **IFAC-PapersOnLine**, v. 49, n. 16, p. 205–208, 2016. Disponível em: <http://dx.doi.org/10.1016/j.ifacol.2016.10.038>.

APÊNDICES

APÊNDICE A. Parâmetros climáticos do local.



——Ciclo ocorrido – – Ciclo programado

10: Período que compreende à coleta do dia 09 fev.
11: Período que compreende à coleta do dia 14 fev.
12: Período que compreende à coleta do dia 20 fev.
18: Período que compreende à coleta do dia 23 mar.

APÊNDICE B. Estágio fenológico, T1 e T2 – M6410 IPRO (Precoce).

	09 fev		14 fev		20 fev		23 mar	
	T1	**T2**	**T1**	**T2**	**T1**	**T2**	**T1**	**T2**
Estágio fenológico	(%)							
R1	0	0	0	0	0	0	0	0
R2	0	0	0	0	0	0	0	0
R3	57,1	73,2	7,1	4,9	0	0	0	0
R4	42,9	26,8	23,8	39,0	0	0	0	0
R5	0	0	69,0	56,1	100	100	0	0
R6	0	0	0	0	0	0	71,7	66,7
R7	0	0	0	0	0	0	28,3	33,3
R8	0	0	0	0	0	0	0	0

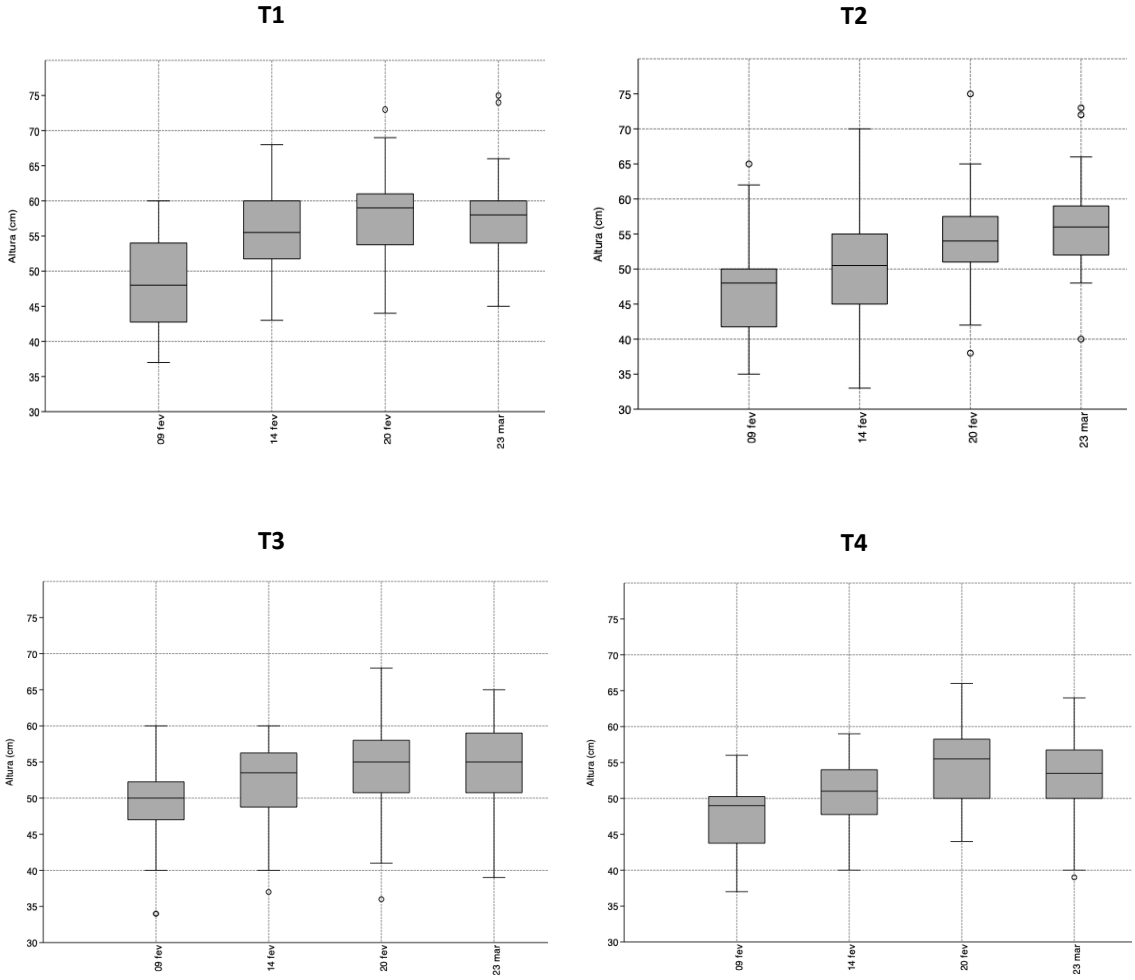
APÊNDICE C. Estágio fenológico, T3 e T4 – M5947 IPRO (Super precoce).

	09 fev		14 fev		20 fev		23 mar	
	T3	**T4**	**T3**	**T4**	**T3**	**T4**	**T3**	**T4**
Estágio fenológico	(%)							
R1	0	0	0	0	0	0	0	0
R2	0	0	0	0	0	0	0	0
R3	40,5	54,8	0	2,4	0	0	0	0
R4	59,5	45,2	28,6	26,2	0	0	0	0
R5	0	0	71,4	71,4	100	100	0	0
R6	0	0	0	0	0	0	63,6	70,8
R7	0	0	0	0	0	0	36,4	29,2
R8	0	0	0	0	0	0	0	0

APÊNDICE D. Componente de produção, mensuradas de acordo com RAS.

			M6410 IPRO		**M5947 IPRO**	
			T1	**T2**	**T3**	**T4**
Vagens por planta			105	120	125	148
Grãos por vagem (%)		**1**	11	14	22	17
		2	37	40	40	43
		3	52	46	38	40
Massa seca (g)			9,2	11,2	11,5	12,6
Retenção em peneira (%)	**Peneiras (mm)**	**5,16**	2	1	1	1
		5,56	8	6	7	5
		5,95	28	22	28	17
		6,35	41	39	39	35
		6,75	0	1	0	0
		7,14	21	31	24	41
Peso de mil grãos (g)			152,4	126,23	121,6	129,56
Taxa de germinação (%)			91,0	94,2	96,0	88,0
Teor de umidade (%)			10,5639	10,7438	10,8308	10,9123

APÊNDICE E. Box-plot da altura das plantas por tratamento.



3. ANNSEED 1.0: PROPOSTA DE TESTE RÁPIDO DE GERMINAÇÃO

RESUMO

A emprego da inteligência artificial em rotinas laboratoriais na área de tecnologia de sementes constitui-se como uma prática ímpar no desenvolvimento de novas técnicas e equipamentos mais modernos, resultando em análises menos morosas e invasivas. Dessa forma, o presente estudo teve como objetivo a elaboração de um novo teste de germinação rápido, eficaz e não-destrutivo através da utilização da inteligência artificial. Para isso, foi utilizado o sensor multiespectral SeedReporter onde foram coletadas as assinaturas espectrais e imagens RGB das sementes. A partir desses dados foi desenvolvida uma rede neural artificial do tipo multicamadas, empregada no processo de predição da germinação de sementes em plântulas normais, anomais ou mortas. Como processo de validação as sementes foram submetidas ao teste de germinação tradicional. No geral, as três versões desenvolvidas de algoritmo foram eficientes na identificação da germinação de sementes indiferente da classe, exibindo uma média de 88% de acerto nas predições e um erro absoluto médio entorno de 0,09. Já para a classe de sementes normais, o acerto foi de entorno de 95%. Em suma, a ferramenta em questão apresentou grande potencial de utilização, sendo possível em um futuro próximo sua implementação às rotinas laboratoriais.

Palavras-chave: Análise de imagem; Multiespectral; Vigor; Aprendizado profundo; Banda do vermelho distante

ABSTRACT

The artificial intelligence use in laboratory routines in seeds' technology area is a unique practice in the development of new techniques and more modern equipment, resulting in less slowly and invasive analysis. In this way, the present study aimed to elaborate a new germination test faster than the traditional one, effective and non-destructive by using artificial intelligence. To achieve that, it was used the SeedReportet multispectral sensor to collect seeds' spectral signatures and their RGB images. With these data were developed a multilayer artificial neural network, used in the prediction process of germination in normal, abnormal, or dead seedlings. As a validation process, the seeds were submitted to the traditional germination test. Overall, the three versions of the algorithm developed were efficient in predicted the seed's germination rate by no particular class, the accuracy average in predictions was 88%, and the mean absolut error was around 0,09. To normal seeds class, the accuracy was around 95%. Briefing, the discussed tool has shown a great use potential, being possible in the near-future its implementation in laboratories routines.

Keywords: Image analysis; Multispectral; Vigour; Deep learning; Far-red band

3.1 INTRODUÇÃO

Ao longo de décadas, a soja [Glycine max (L.) Merril.] tem sido a mais importante commodity negociada do Brasil (COELHO et al., 2018). Com uma produção de 135,4 milhões de toneladas na safra 2020/21, o Brasil tem se estabelecido como o maior produtor e exportador da leguminosa no mundo (CONAB, 2021). Atualmente, o cultivo da leguminosa tem se estendido por quase todo território nacional, ocupando cerca de 36 milhões de hectares (APROSOJA, 2019). Essa vasta distribuição pelo território brasileiro, faz com que a produção de sementes transcorra nas mais diversas condições ambientais e financeiras. Neste contexto, visando garantir a qualidade dos lotes comercializados, foram estabelecidos padrões de identidade e qualidade das sementes, a partir da instrução normativa 45/2013 (BRASIL, 2013). Além disso, todas as análises de sementes foram padronizadas através de rígidos protocolos estabelecidos pelo Ministério da Agricultura, Pecuária e Abastecimento (BRASIL, 2009), sendo um dos mais importantes teste de controle de qualidade, o teste de germinação.

O teste de germinação tem como objetivo determinar o potencial máximo de germinação de um lote de sementes, através da identificação e quantificação dos indivíduos que possuem a capacidade de emergir em uma plântula normal, em condições ótimas de temperatura, iluminação e água (BRASIL, 2009). Com os resultados em mãos, os lotes podem ser negociados no mercado de commodities, armazenados em silos e/ou exportados, ressaltando que, resultados abaixo do mínimo tolerado pelas Regras de Análise de Sementes (RAS) desqualificam o produto de semente para grão ocorrendo a desvalorização do lote (BRASIL, 2009; APROSOJA, 2019; CONAB, 2021). Nesse contexto, a estimativa do potencial fisiológico de sementes é rotineiramente avaliada por teste de germinação em conjunto com outros testes, essenciais para apontar diferenças relacionadas ao desempenho dos lotes de sementes (MARCOS FILHO et al., 2009). Contudo, os métodos de avaliação da germinação de sementes são processos considerados morosos, destrutivos e subjetivos (HONG et al, 2015). Desse modo, novas metodologias estão sendo consideradas mais efetivas, como a análise de imagens de sementes aliadas à inteligência artificial.

Inteligência artificial (IA) é um conjunto de ferramentas que são capazes de permitir que sistemas e computadores possam ver, aprender, ententer e decidir sobre situações as que estão expostas (ABDAR et al., 2021). O emprego dessas ferramentas provindas da IA tem se tornado cada vez mais comum na área de tecnologia de sementes. Segundo análises do SCOPUS (2021), os termos “agricultura, sementes, imagem e IA” foram registrados juntos pela primeira vez em 1988, sendo o primeiro trabalho publicado com soja em 1989, atualmente há 1.371 registros de artigos contendo os termos juntos. Ainda segundo as análises, a partir de 2006 houve um aumento exponencial das pesquisas, especialmente em três áreas “Agricultura e Ciências Biológicas” (21,8%), “Engenharia” (16,7%) e “Ciência da Computação” (16,6%), sendo as pesquisas lideradas pela China seguida pelos EUA e India (SCOPUS, 2021). Atualmente, o desenvolvimento de novas técnicas para análise de sementes vem ocorrendo em duas etapas: *computer vision* e *machine learning* (LURSTWUT; PORNPANOMCHAI, 2017).

A etapa de *computer vision*, ou visão computacional, permite que sistemas e computadores através da captura de imagens e/ou vídeos vejam, observem e compreendam de maneira automatizada o meio, podendo dessa forma, estabelecer padrões ou tomar decisões previamente definidas (IBM, 2021). Essa captura pode ocorrer através dos mais diversos sensores de coleta. Após a coleta o produto capturado precisa passar por pós-processamento para que as informações sejam extraídas. Neste contexto, tem sido utilizado o *software* ImageJ, desenvolvido em meados de 2012 por Wayne Rasband para a área da saúde. Atualmente o software é considerado uma eficiente ferramenta de domínio público no auxílio de processamento e análises das imagens tanto no segmento da saúde, como também na agricultura (FERREIRA; RASBAND, 2012).

Posteriormente à etapa *computer vision*, ocorre a etapa de *machine learning*, ou aprendizado de máquina. Nessa etapa, a máquina passa a aprender de maneira iterativa através de treinos, onde uma série de respostas esperadas é gerada por meio de associações de diferentes dados, extraídos da etapa anterior, gradualmente a mesma passa a tomar melhores decisões baseadas na experiência prévia de treino (IBM, 2021). Para esse processo de aprendizado e treino tem sido comumente utilizado o *software* WEKA (*Waikato Environment for Knowledge Analysis*), desenvolvido em 1993 pela University of Waikato. O programa possui uma plataforma intuitiva, de domínio público, construída para aplicação de técnicas de *machine learning* via implementação de algoritmos diversos (BOUCKAERT et al., 2012).

A união de ambas as etapas à área de tecnologia de sementes, tem se mostrado uma poderosa aliada nas análises laboratorias e, vem exibindo resultados promissores. SKRUBEJ et al. (2016) ao construírem um sistema visão computacional baseado em processamento de imagens e redes neurais artificiais (RNA), com os *softwares* ImageJ e WEKA, classificaram corretamente 95,44% das sementes germinadas de tomateiro (*Solanum lycopersicum* L.). ANDRADE et al. (2016) ao usarem computação visual e máquina de aprendizado melhoraram o beneficiamento da castanha brasileira, distinguindo corretamente castanhas intactas e quebradas das boas e estragadas. LURSTWUT e PORNPANOMCHAI (2017) desenvolveu um sistema de avaliação de sementes de arroz para predição da germinação utilizando imagens e técnica de redes neurais artificiais, a taxa de precisão do sistema foi de 7,66% falso positivo e 5,42% falso negativo para a germinação. MONTEIRO et al. (2020) desenvolveu uma metodologia capaz de identificar através do processamento de imagem de sementes de soja indivíduos com danos causados por variações de temperatura e umidade, sendo possível identificar o tipo de dano (semente esverdeada ou enrugada) e dentro das enrugadas o nível das rugas.

Portanto, o objetivo deste presente capítulo foi elaborar um novo teste de germinação rápido e eficaz, baseado em análise de imagens e inteligência artificial.

3.2 MATERIAL E MÉTODOS

O presente trabalho foi conduzido no Laboratório de Análise de Sementes e Laboratório de Imagens, do Departamento de Produção Vegetal – USP/ESALQ, segundo orientações do Ministério da Agricultura, Pecuária e Abastecimento (BRASIL, 2009). Foram utilizadas 800 sementes de soja (Glycine max L.), cultivar M5947 IPRO safra 2018/2019, os indivíduos foram selecionados aleatoriamente.

A metodologia elaborada para o presente estudo divide-se em três etapas: aquisição de dados, extração de características físicas e espectrais e a elaboração do algoritmo. Paralelamente, a esses processos foi executada a contraprova da eficiência do algoritmo, através do teste de germinação tradicional seguindo os protocolos estabelecidos pelo Ministério da Agricultura nas RAS.

3.2.1 Etapa 1: aquisição de dados

No processo de aquisição de dados, foram coletadas a imagens RGB e a assinatura espectral de cada uma das sementes. Ambas as coletas ocorreram concomitantemente dentro do mesmo equipamento. O equipamento utilizado nesse processo foi o sensor multiespectral SeedReporter precursor do PlantExplorer, ambos fabricados pela empresa PhenoVation™. O sensor captura imagens multiespectrais em 14 bits, através da emissão de subpulsos dos LEDs em uma área de 20x20cm^2. As amostras adentram a câmara de isolamento em placas pretas de PVC, garantindo assim que a única fonte de luz incidente sobre os indivíduos fosse proveniente do sensor, conforme a figura abaixo.

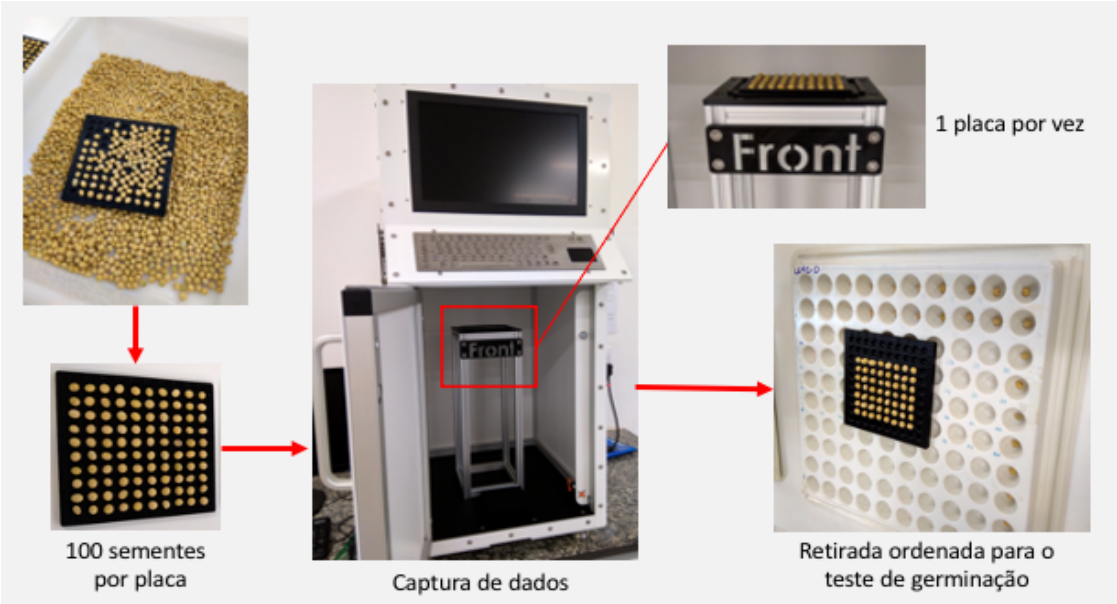


Figura 1. Etapa de aquisição de dados.

3.2.2 Teste de germinação

Após a captura de dados, as sementes foram submetidas imediatamente ao teste de germinação tradicional seguindo os protocolos da RAS (BRASIL, 2009), todos os indivíduos foram numerados um-a-um e distribuídos sobre 3 folhas de papel Germitest®, previamente umedecido em água destilada, na proporção de massa de 2,5:1 (água:papel). Os rolos com as sementes foram levados ao germinador à 20ºC e as contagens realizadas no sétimo dia após a instalação, conforme a Figura 2.

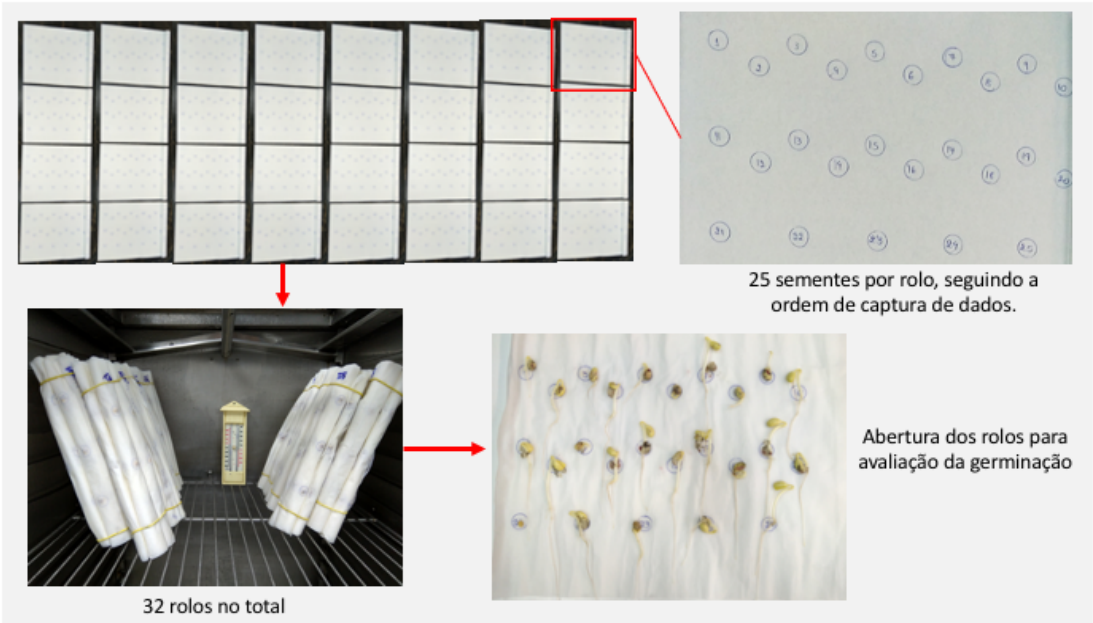


Figura 2. Teste de germinação tradicional executado.

Visando atenuar os erros de classificação entre as categorias pré-determinadas pelo RAS (plântulas normais, anormais e mortas), todas às 800 avaliações foram acompanhadas por uma especialista técnica em fisiologia de sementes do Departamento de Produção Vegetal – USP/ESALQ. Todos os protocolos estabelecidos pelo Ministério da Agricultura, foram seguidos de maneira rígida, obedecendo à ordem de captura de dados e entrada do rolo de sementes no germinador. Por conseguinte, reduzindo qualquer fonte de erro externa ao algoritmo no processo de validação do mesmo. Uma vez que, o teste de germinação foi utilizado no processo de desenvolvimento da rede neural e de validação final da mesma. Segue abaixo o padrão de indivíduos por classe, estabelecido durante as avaliações.

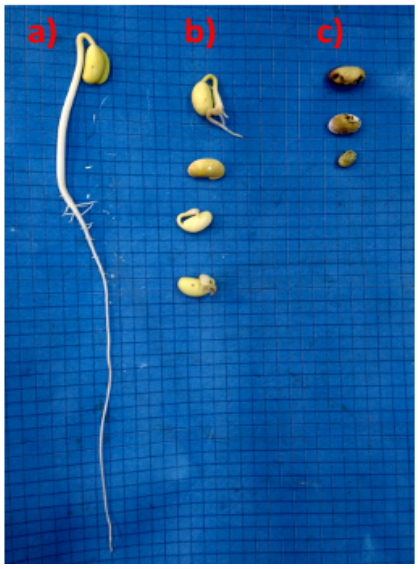


Figura 3. Exemplo de indivíduos de cada classe encontrados no lote utilizado para estudo. a) Normal, b) Anormal, c) Morta.

3.2.3 Etapa 2: extração das características físicas e espectrais

Após o teste de germinação foi feito o processo de extração de características físicas e espectrais, através de pós-processamento de dados elaborado pelo fabricante do SeedReporter. Primeiramente, foram extraídas as características espectrais. Dentro do software as sementes receberam identificação e, para cada indivíduo fotografado obteve as seguintes informações extraídas:

- Imagem RGB;
- Bandas azul, verde, vermelho, vermelho distante, infravermelho próximo;
- NDVI (Índice de vegetação com diferença normalizada);
- Chl (Correlacionado com a quantidade de clorofila presente no indivíduo);
- Hue (Matiz);
- Saturação.

Em posse das imagens RGB, através do software ImageJ foram extraídas as características físicas das sementes. As imagens foram submetidas ao processo de segmentação de imagens aplicando a técnica de limiarização (*thresholding*), conforme exposto por alguns autores (ANDRADE et al., 2016; SHARMA; ABROL, 2016; SKRUBEJ; ROZMAN; STAJNKO, 2016; LURSTWUT; PORNPANOMCHAI, 2017). Essa técnica que também é conhecida como a binarização de uma imagem, é utilizada para separa objetos através dos níveis de cinza, ou seja, duas regiões (objeto e o fundo) são desagregadas a partir de um limiar, onde os valores dos pixels menores são mapeados tendendo a 0 e os valores maiores tendendo a 1 (em imagens de 8 bits mapeados tendedo de 0 a 255) (RIDLER; CALVARD, 1978; FERREIRA; RASBAND, 2012; MARIA; TOMMASELLI, 2018). A principal ferramenta de avaliação da qualidade da limiarização executada é o histograma de frequência, verifica-se que uma “ótima limiarização” deverá exibir dois picos distintos havendo um vale demarcado entre os picos, cada um deles representará uma das classes (objeto e o fundo) (MARQUES FILHO; VIEIRA NETO, 1999).

Para este estudo o processo de limiarização da imagem RBG em um primeiro momento não se mostrou eficiente, não havendo um vale demarcado e apresentando uma ineficiência na seleção de indivíduos, o que prejudicaria a coleta das características físicas de cada uma das sementes. Dessa forma, a imagem RGB teve suas as bandas separadas em um procedimento entitulado de *split channels*, sendo o resultado submetido novamente ao processo de segmentação de imagens aplicando a técnica de limiarização (SHARMA; ABROL, 2016). Logo, foi notado que a banda do vermelho e azul foram as que mais se aproximaram de uma “ótima limiarização”, exibindo um histograma com picos e vale demarcado, além de uma seleção de indivíduos bastante eficiente. Contudo, a banda do azul foi a que apresentou uma melhor eficiência na seleção de indivíduos de maneira automática, não necessitando de ajustes finos posteriormente, neste contexto, optou-se pela utilização da banda do azul como fonte de coleta de dados para o algoritmo. Segue na Figura 4, o processo de segmentação de imagens aplicando a técnica de limiarização tanto na imagem RGB quando nas imagens com as bandas separadas.

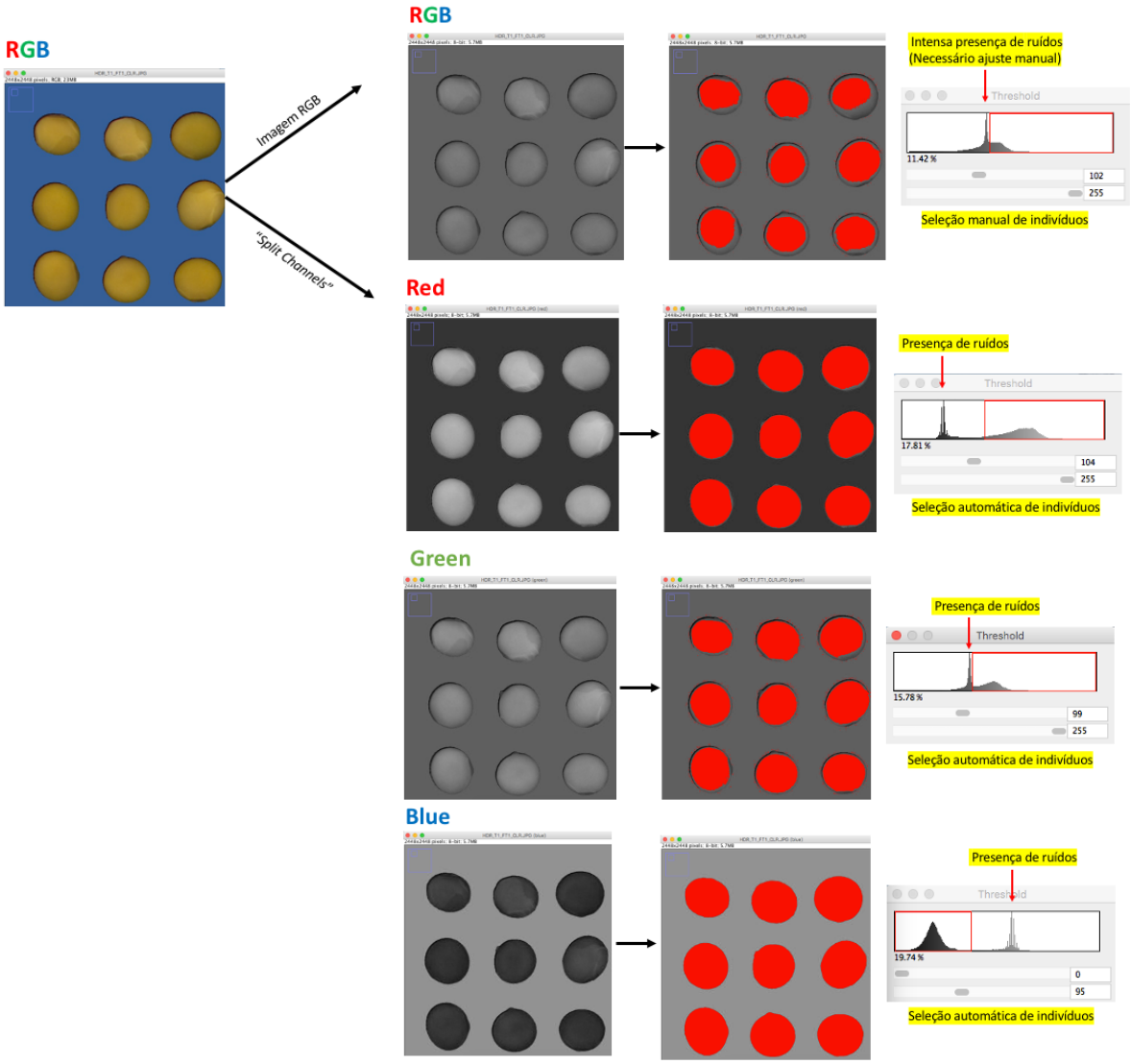


Figura 4. Processo de segmentação das imagens através da limiarização no ImageJ.

Após a seleção dos indivíduos, foram extraídas nove características de cada uma das sementes, sendo a escolha das mesmas feita após análise de trabalhos semelhantes (DUCOURNAU et al., 2004; ANDRADE et al., 2016; SKRUBEJ; ROZMAN; STAJNKO, 2016; LURSTWUT; PORNPANOMCHAI, 2017; RAJU AHMED et al., 2020). Na tabela 1 a seguir, segue os parâmetros físicos utilizados e suas descrições.

Tabela 1. Parâmetros físicos extraídos na etapa 2.

Parâmetro	Descrição
Area	Área ocupada pelo indivíduo em mm^2
Mean	Média dos valores de cinza que compõem o indivíduo selecionado
StdDev	Desvio padrão dos valores de cinza que compõem o indivíduo selecionado
Perim.	Perímetro da área ocupada pelo individuo em mm
Circ.	Avalia o quão circular é um indivíduo a partir do perímetro, podendo variar entre 0 (linha) e 1 (círculo perfeito)
Feret	Maior comprimento presente no indivíduo em mm
AR	Proporção de tela, uma divisão entre a maior largura e o maior comprimento
Round	Avalia o quão circular é um indivíduo a partir diâmetro máximo, podendo variar entre 0 (linha) e 1 (círculo perfeito)
Solidity	Avalia o quão sólido é o indivíduo a partir da divisão entre a área pela área convexa

Todas as informações coletadas foram exportadas como arquivo csv para posteriormente ingressarem na construção do algoritmo.

3.2.4 Etapa 3: elaboração do algoritmo de predição

Algoritmo é definido como sendo uma sequência de regras estabelecidas através de um raciocínio lógico, visando a execução de operações matemáticas capazes de solucionar determinado problema. O principal objetivo dos algoritmos é a otimização de processos, garantindo assim agilidade, eficiência e padronização. Os algoritmos segundo Alan Turing, instituidor da computação moderna, precisam obrigatoriamente seguir 3 regras fundamentais:

1. Não possuir ações ambíguas;
2. Suas ações devem seguir uma execução lógica;
3. Possuir início e fim.

Neste contexto, buscando elaborar um algoritmo que pudesse receber vários tipos de informações e operacionaliza-las simultaneamente no mesmo ambiente, optou-se pela elaboração de uma rede neural artificial (RNA) do tipo multicamadas de Perceptrons (MLP) via retropropagação (*Backpropagation*). As redes neurais artificiais são sistemas de aprendizado interconectados semelhantes ao funcionamento da rede neural presente no cérebro humano e, os perceptrons são estruturas de processamento de dados análogas aos neurônios, ou seja, através dessa estrutura o algoritmo pode estabelecer padrões mesmo nos dados mais complexos e não-lineares (PAULI, 2020; ABDAR et al., 2021; BARANAUSKAS, J. A., 2021). Em suma, essa estrutura pode permitir que o algoritmo tenha um desempenho paralelo ao humano ao longo dos anos, havendo espaço para melhorias e aprimoramentos de acordo com as necessidades advindas das futuras demandas. Dessa forma, o algoritmo de predição elaborado no presente capítulo, recebeu da presente autora deste trabalho o nome de **ANNSEED** (**A**rtificial **N**eural **N**etwork for **Seed**), sendo seu objetivo predizer se uma semente de soja irá germinar ou não, utilizando como base dados fornecidos por imagens.

3.2.5 Treino e teste do ANNSEED

Para a construção do algoritmo as amostras foram separadas atráves do método de amostragem *houldout*, onde são estabelecidos dois grupos excludentes com porcentagens fixas de amostras, o treino "*p"* e validação "*1-p"* (SCHNEIDER, 1997; OMARY; MTENZI, 2010; DOBBIN; SIMON, 2011; PEREIRA; CENTENO, 2017; ABDAR et al., 2021; BARANAUSKAS, J. A., 2021). Essa metodologia garantiu que não houvesse uma releitura das amostras fora do grupo ao qual pertenciam, ou seja, uma amostra selecionada para treino não poderia integrar o grupo de validação e o contrário se fez verdadeiro.

Seguindo a metodologia de amostragem, primeiramente os grupos foram separados de acordo com a classes de germinação (Normal, Anormal e Morta), conforme a figura 3 apresentada anteriormente. Em um segundo momento, optou-se por separar metade dos indivíduos de cada classe para o grupo de treino e a outra metade para validação (ou seja, porcentagem fixa de 50% para cada grupo). Por fim, a escolha das amostras que incorporaram cada um dos grupos foi feita através do *software* WEKA utilizando o filtro *supervised instance resample*, usado para reamostrar grupos através de classes nominais podendo repor a amostra ou não (BOUCKAERT et al., 2012) para o presente estudo não foram admitidas amostras repetidas para montagem do grupo de treino e validação.

Para os testes de validação o grupo composto por 400 indivíduos selecionados anteriormente, passou por 10 sorteios onde foram selecionadas sementes para compor grupos teste podendo haver repetição e ausência de classe. Dessa forma, simulando a amostragem realizada atualmente pela RAS (BRASIL, 2009).

3.2.6 Rede neural artificial multilayer perceptron (RNA-MLP)

Para o presente capítulo foi escolhida como principal base de elaboração do ANNSEED a estrutura RNA-MLP, devido a sua facilidade de adaptação, processamento mais leve e, por ser altamente difundida sendo empregada desde 1986 nos estudos mais diversos (PAULI, 2020; ABDAR et al., 2021). A RNA-MLP é formada por uma série de nós que em conjunto formam as camadas, existem obrigatoriamente no mínimo três tipos de camada: camada de entrada (*input layer*), camadas ocultas (*hidden layers*) e a camada de saída (*output layer*).

Esses arranjos são interligados por meio das pontes, estruturas de interação, sendo por meio delas propagado o conhecimento adquirido em cada um dos nós pertencentes as camadas ocultas (BARANAUSKAS, J. A., 2021). Essa propagação é um processo iterativo que pode ser executado inúmeras vezes, sendo parte essencial ao aprendizado de máquina. Cada vez que essa iteração é executada, maior é o aprendizado adquirido em cada nó, esse processo é intitulado de retropropagação de aprendizado (*learning backpropagation*) e, garante que as pontes sejam rebalanceadas adequadamente minimizando os erros de classificação. Contudo, existe um limite para a taxa de aprendizado, ou seja, maiores quantidades de repetições obterão o mesmo nível de aprendizado, estabelecendo-se assim o "limite máximo de aprendizado" (SKRUBEJ; ROZMAN; STAJNKO, 2016).

Os nós das camadas ocultas adquirem conhecimento por meio da função de ativação escolhida para uso no algoritmo, essa ativação pode ser por degrau, identidade, sigmoide logística ou tangente hiperbólica (PAULI, 2020). Para este presente estudo optou-se pela função sigmoide logística, pois seus parâmentros

permitiram um aprendizado com maior grau detalhamento e entendimento sobre a germinação de sementes, o que por consequência resultou em um melhor balanceamento dos pesos presentes nas pontes. Segue abaixo a representação gráfica da RNA-MLP.

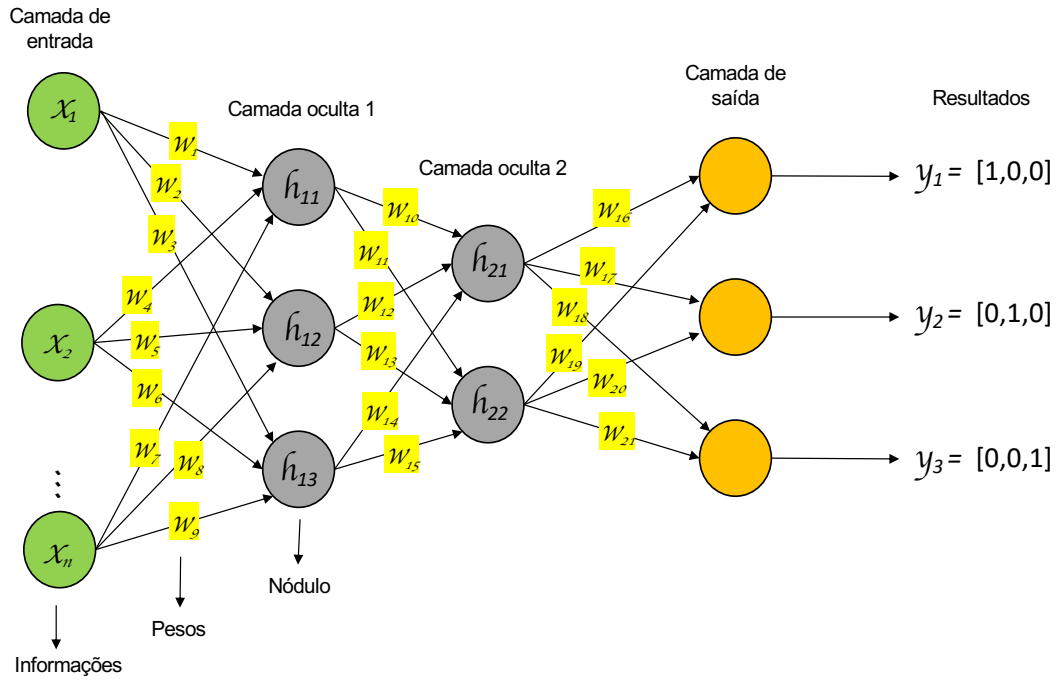


Figura 5. Representação gráfica de uma RNA-MLP.

Todas as etapas compreendidas na metodologia seguem ilustradas abaixo.

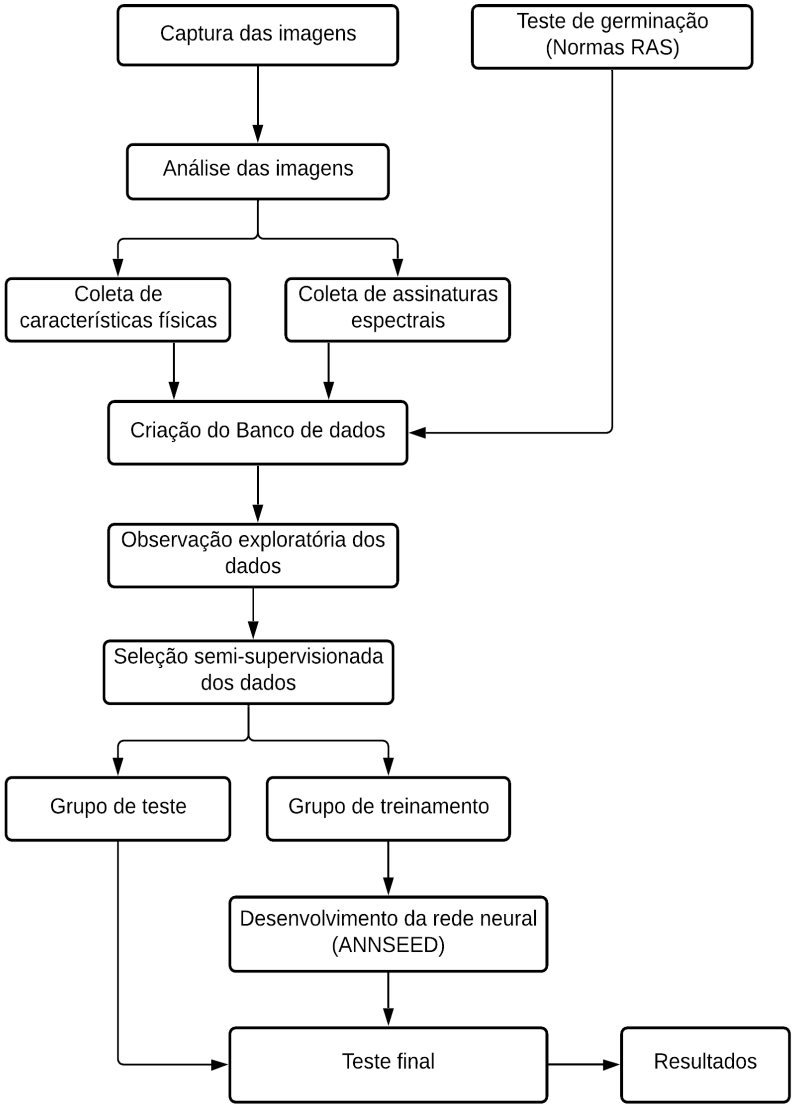


Figura 6. Diagrama geral dos procedimentos aplicados na elaboração do ANNSEED.

3.3 RESULTADOS E DISCUSSÃO

Como resultado do presente estudo, foram elaboradas três versões do ANNSEED sendo elas constituídas da seguinte forma: somatório de *inputs* de dados espectrais e físicos (Versão A), *input* apenas de dados espectrais (Versão B) e *input* apenas de dados físicos (Versão C). Uma vez, constituídas as três versões, as mesmas foram submetidas ao processo de validação, onde foi avaliada a e viabilidade de cada uma das versões. Segue a representação gráfica RNA-MLP do ANNSEED em cada uma das versões.

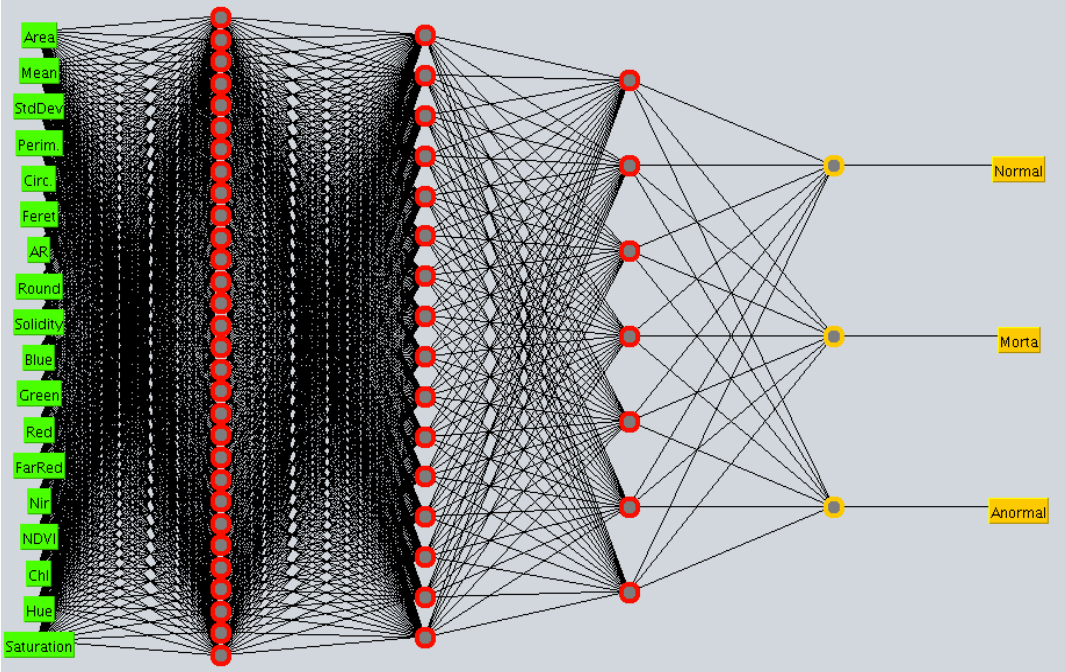


Figura 7. Versão A do ANNSEED 1.0, união dos inputs espectrais e físicos.

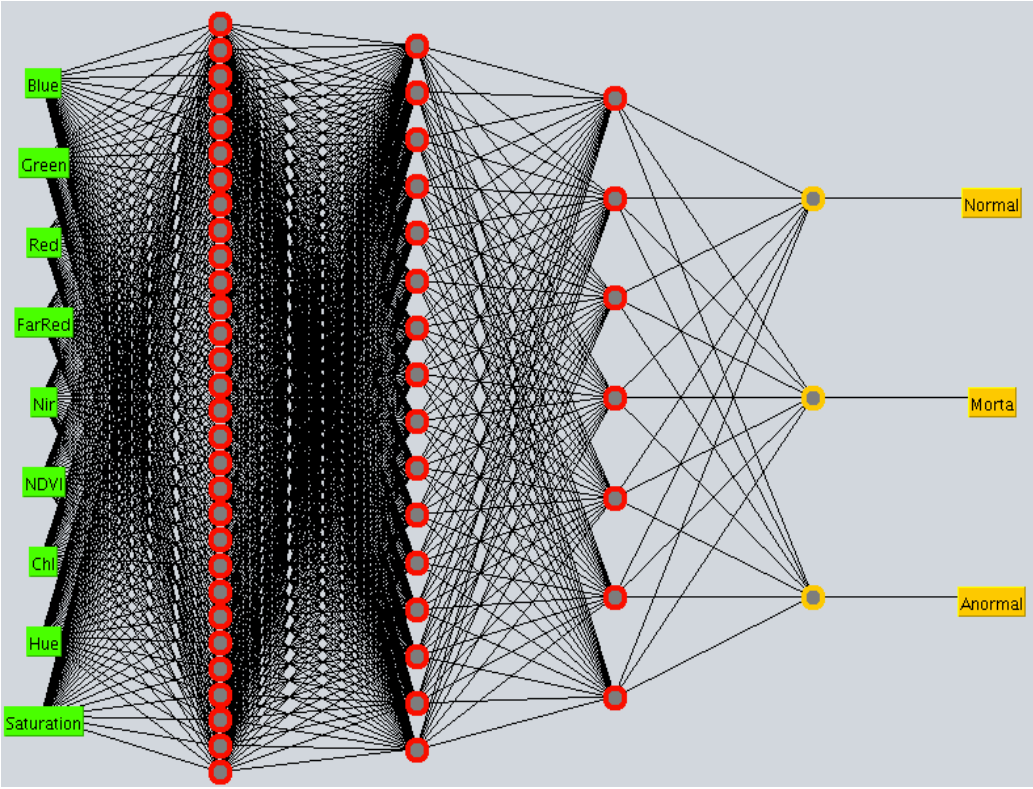


Figura 8. Versão B do ANNSEED 1.0, inputs espectrais.

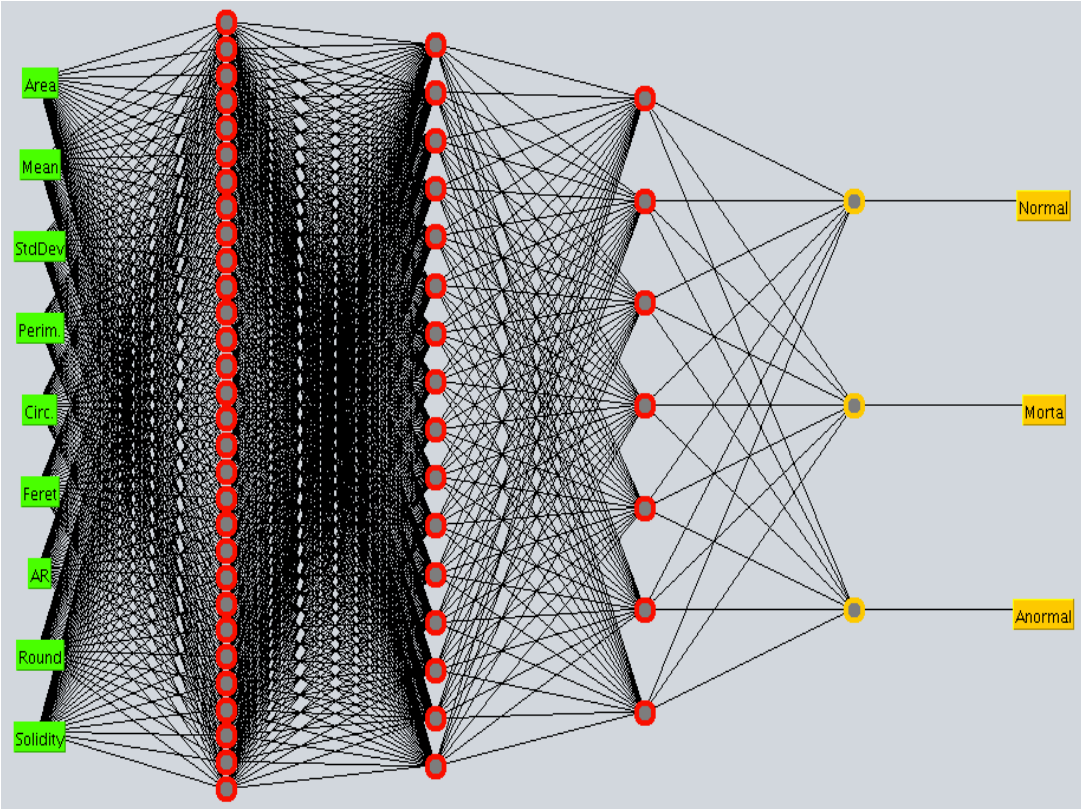


Figura 9. Versão C do ANNSEED 1.0, inputs físicos.

Uma vez, elaborado o algoritmo de predição, foi realizado o teste de germinação segundo protocolos da RAS (BRASIL, 2009), os resultados mostraram que o lote constituído pelas 800 amostras, possuía em sua composição: 92% de plântulas normais, 6,75% anormais e 1,25% de sementes mortas. Dessa forma, o lote foi classificado comercialmente como sendo de sementes (BRASIL, 2009; APROSOJA, 2019; CONAB, 2021). Ressaltando ainda, que no total o teste de germinação levou 7 dias para ficar pronto, podendo se prolongar por ainda mais 3 dias caso fosse necessário, uma vez que, lotes de soja precisam de no mínimo 5 dias no germinador e no máximo 10 dias, ficando a critério do avaliador quando encerrar o mesmo dentro do período exposto (BRASIL, 2009). Outro ponto relevante, o teste de germinação tradicional é essencialmente um teste destrutivo, inviabilizando qualquer contraprova. Segue a composição do lote de soja utilizado, bem como, a composição do grupo de treino utilizados para a elaboração do ANNSEED e do grupo de validação utilizados na avaliação final do algoritmo.

Tabela 2. Composição das amostras para o grupo de treino e grupo de validação.

Classe de germinação	Total de indivíduos	Proporção biológica do lote (%)	Grupo de Treino	Grupo de Validação
Normal	736	92,00	368	368
Anormal	54	6,75	27	27
Morta	10	1,25	5	5

Como segundo resultado, as versões elaboradas do ANNSEED tiveram que classificar diversos tamanhos de lotes. Nessa classificação, foi observado que de maneira geral entre lotes compostos de 40 a 400 indivíduos

o percentual de erro do ANNSEED foi abaixo de 18%, sendo 40 indivíduos o percentual mais alto (17,5%) e 360 indivíduos o percentual mais baixo (8,9%). Dentre todas as versões, a que se mostrou mais estável foi a versão C, exibindo menor média, desvio padrão e amplitude, seguida pela versão B e versão A. Considerando que a versão A foi composta pelo somatório dos *inputs* da versão B e versão C, pode-se dizer que as características juntas geraram uma maior confusão ao ANNSEED durante as classificações. Segue abaixo, a plotagem do percentual de sementes classificadas erroneamente e tabela com os valores na integra.

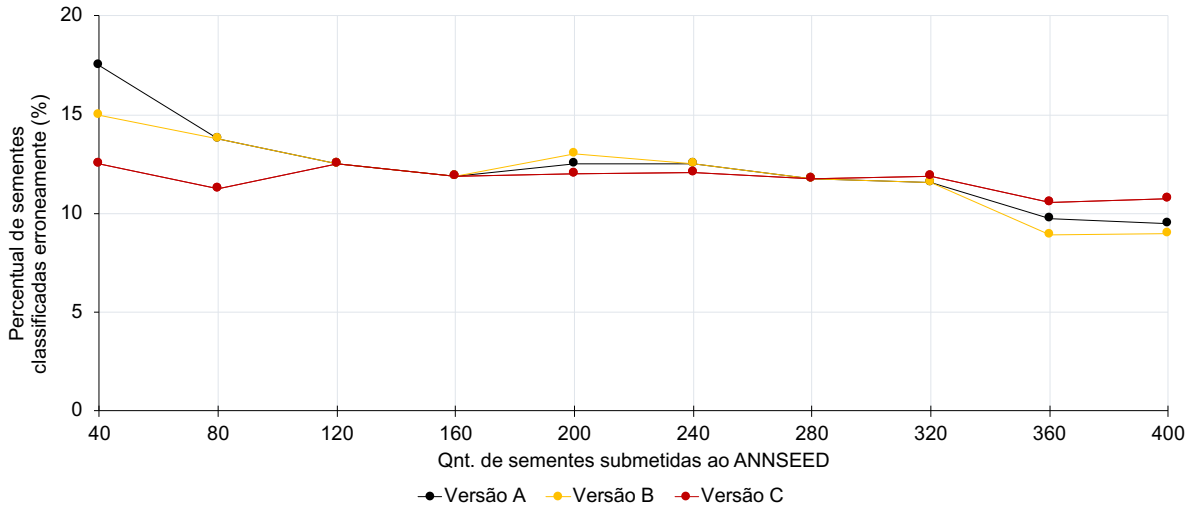


Figura 10. Gráfico do percentual de sementes classificadas erroneamente.

Tabela 3. Percentual de sementes classificadas erroneamente na avaliação final.

Percentual de sementes classificadas erroneamente (%)			
Qnt. amostras	**Versão A***	**Versão B****	**Versão C*****
40	17,5	15,0	12,5
80	13,8	13,8	11,3
120	12,5	12,5	12,5
160	11,9	11,9	11,9
200	12,5	13,0	12,0
240	12,5	12,5	12,1
280	11,8	11,8	11,8
320	11,6	11,6	11,9
360	9,7	8,9	10,6
400	9,5	9,0	10,8
Média	12,3	12,0	11,7
Amplitude	2,1	1,8	0,6
Desvio padrão	8	6,1	1,9

*Versão A: *Input* de dados espectrais e físicos **Versão B: *Input* de dados espectrais ***Versão C: *Input* de dados físicos

Como terceiro resultado, foi avaliado o desempenho do algoritmo através do cálculo do erro absoluto médio (MAE) entre a classificação observada e a predita, onde tem-se que quanto menor é o valor do erro, melhor ajustado aos parâmetros o algoritmo está. De maneira geral, o ANNSEED em suas três versões apresentou resultados promissores, sendo em todos os testes abaixo de 0,12 e exibindo uma estabilidade no gráfico. Neste contexto, observou que a versão C foi novamente a que exibiu a menor média, desvio padrão e amplitude, seguida pela versão A e versão C. Ressaltando que todas as três versões possuem a mesma arquitetura RNA-MLP

e utilizam mesma função de ativação, esse resultado demonstrou que os *inputs* utilizados na versão B não ficaram tão bem ajustados quanto os da versão C ao ANNSEED. Outro ponto relevante, foi que a versão A apresentou resultado semelhantes ao da versão C no que diz respeito ao MAE. Isso mostrou o quão bem ajustados os *inputs* da versão C foram a RNA-MLP, podendo inferir que eles foram responsáveis diretamente pelos resultados exibidos na versão A. Segue abaixo, a plotagem do erro absoluto médio e tabela com os valores na íntegra.

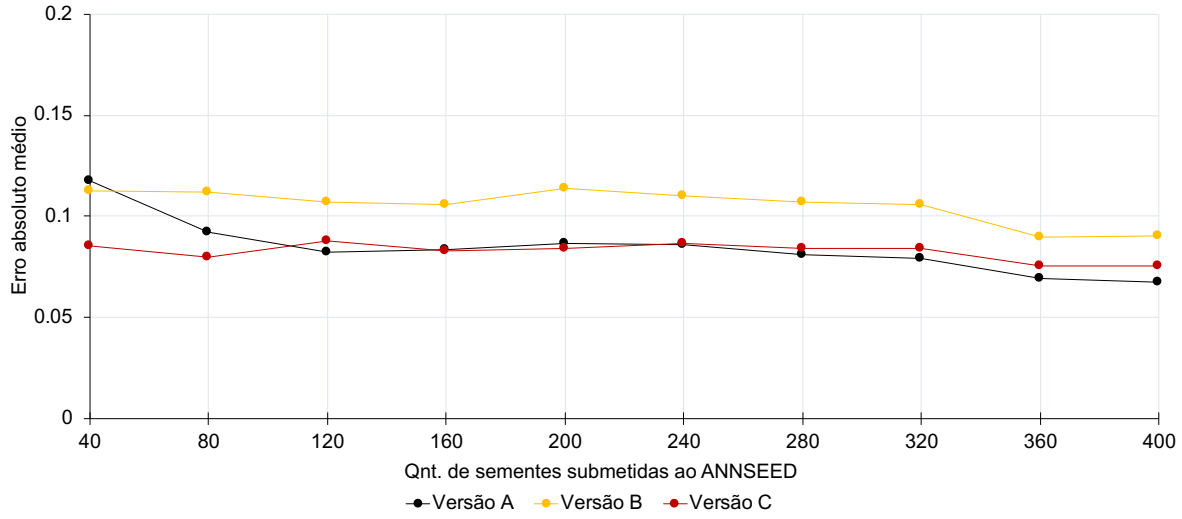


Figura 11. Gráfico do erro médio absoluto (MAE).

Tabela 4. Erro absoluto médio (MAE).

MAE			
Qnt. amostras	**Versão A***	**Versão B****	**Versão C*****
40	0,118	0,113	0,086
80	0,092	0,112	0,080
120	0,082	0,107	0,088
160	0,084	0,106	0,083
200	0,087	0,114	0,084
240	0,086	0,110	0,087
280	0,081	0,107	0,084
320	0,080	0,106	0,084
360	0,069	0,090	0,075
400	0,067	0,090	0,076
Média	0,085	0,105	0,083
Amplitude	0,013	0,008	0,004
Desvio padrão	0,050	0,024	0,013

*Versão A: *Input* de dados espectrais e físicos **Versão B: *Input* de dados espectrais ***Versão C: *Input* de dados físicos

Como quarto resultado, foi mensurado o índice Kappa, esse indicador avalia o grau de concordância entre os dados, ou seja, a precisão e confiabilidade nas classificações. Quanto mais próximo de 1 for o valor, mais alta será confiança na classificação (Classificação dos valores de Kappa: Menor que 0 – Ausência de confiança; 0 a 0,19 – Confiança pobre; 0,20 a 0,39 – Confiança leve; 0,40 a 0,59 – Confiança moderada; 0,60 a 0,79 – Confiança substantiva; 0,80 a 1 – Confiança quase perfeita). Neste contexto, observou que a versão A foi a que apresentou maior precisão e confiabilidade nas classificações, seguida pela versão C e B. Outro fato relevante foi a ausência

de confiança em lotes iguais ou menores que 40 para versão A, iguais ou menores que 120 para versão C e, iguais ou menores que 160 para versão C. Segue abaixo, a plotagem do índice Kappa e tabela com os valores na íntegra.

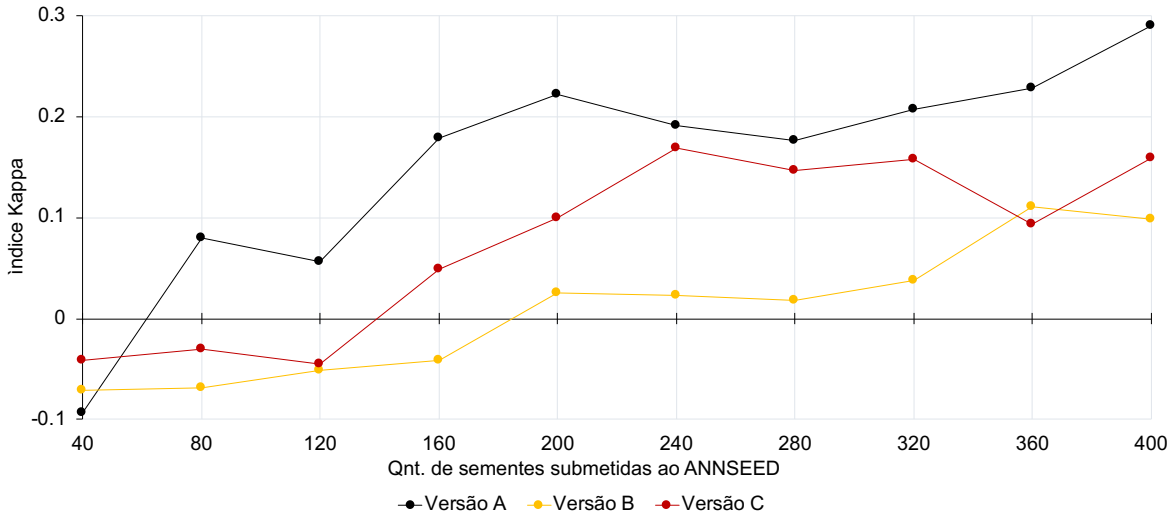


Figura 12. Gráfico do índice Kappa.

Tabela 5. Índice Kappa.

	Kappa		
Qnt. amostras	**Versão A***	**Versão B****	**Versão C*****
40	-0,09	-0,07	-0,04
80	0,08	-0,07	-0,03
120	0,06	-0,05	-0,04
160	0,18	-0,04	0,05
200	0,22	0,02	0,10
240	0,19	0,02	0,17
280	0,18	0,02	0,15
320	0,21	0,04	0,16
360	0,23	0,11	0,09
400	0,29	0,10	0,16

*Versão A: *Input* de dados espectrais e físicos **Versão B: *Input* de dados espectrais ***Versão C: *Input* de dados físicos

Como último resultado, foi capturada a matriz de confusão para cada um dos testes em cada uma das versões. Esses resultados permitiram visualizar o desempenho de cada uma das versões do ANNSEED com relação a classificação feita por ele. Nas linhas foram representadas as classificações verdadeiras, já nas colunas as classificações preditas. Nesse contexto, foi possível perceber uma tendência nas três versões em classificar erroneamente as sementes anormais. Além disso, foi possível perceber que mesmo com uma quantidade reduzida de amostras para classe de sementes mortas em sua biblioteca (5 inidvíduos), as três versões do ANNSEED conseguiram classificar entregar resultados promissores. Em suma, as três versões apresentaram poucos erros nas classificações de sementes normais, exibindo altos percentuais de acerto nessa classe (cerca de 95%). Fato relevante, uma vez que essa classe é a mais vislumbrada comercialmente. Afinal, é através dela que os lotes são classificados comercialmente. Segue as matrizes de confusão resultantes de cada um dos testes em cada uma das versões.

Tabela 6. Matrizes de correlação resultantes.

Versão A*	Versão B**	Versão C***
Avaliação final: 40 amostras		
a b c <-- classified as 33 0 3 | a = Normal 0 0 0 | b = Morta 4 0 0 | c = Anormal	a b c <-- classified as 34 0 2 | a = Normal 0 0 0 | b = Morta 4 0 0 | c = Anormal	a b c <-- classified as 35 0 1 | a = Normal 0 0 0 | b = Morta 4 0 0 | c = Anormal
Avaliação final: 80 amostras		
a b c <-- classified as 68 0 5 | a = Normal 0 0 0 | b = Morta 6 0 1 | c = Anormal	a b c <-- classified as 69 0 4 | a = Normal 0 0 0 | b = Morta 7 0 0 | c = Anormal	a b c <-- classified as 71 1 1 | a = Normal 0 0 0 | b = Morta 7 0 0 | c = Anormal
Avaliação final: 120 amostras		
a b c <-- classified as 104 0 5 | a = Normal 0 0 0 | b = Morta 10 0 1 | c = Anormal	a b c <-- classified as 105 0 4 | a = Normal 0 0 0 | b = Morta 11 0 0 | c = Anormal	a b c <-- classified as 105 1 3 | a = Normal 0 0 0 | b = Morta 11 0 0 | c = Anormal
Avaliação final: 160 amostras		
a b c <-- classified as 138 0 7 | a = Normal 0 0 0 | b = Morta 12 0 3 | c = Anormal	a b c <-- classified as 141 0 4 | a = Normal 0 0 0 | b = Morta 15 0 0 | c = Anormal	a b c <-- classified as 140 1 4 | a = Normal 0 0 0 | b = Morta 14 0 1 | c = Anormal
Avaliação final: 200 amostras		
a b c <-- classified as 170 0 9 | a = Normal 1 0 0 | b = Morta 15 0 5 | c = Anormal	a b c <-- classified as 173 0 6 | a = Normal 0 1 0 | b = Morta 20 0 0 | c = Anormal	a b c <-- classified as 174 1 4 | a = Normal 1 0 0 | b = Morta 18 0 2 | c = Anormal
Avaliação final: 240 amostras		
a b c <-- classified as 205 0 10 | a = Normal 3 0 0 | b = Morta 17 0 5 | c = Anormal	a b c <-- classified as 209 0 6 | a = Normal 2 1 0 | b = Morta 22 0 0 | c = Anormal	a b c <-- classified as 207 1 7 | a = Normal 1 2 0 | b = Morta 20 0 2 | c = Anormal
Avaliação final: 280 amostras		
a b c <-- classified as 242 0 11 | a = Normal 3 0 0 | b = Morta 19 0 5 | c = Anormal	a b c <-- classified as 246 0 7 | a = Normal 2 1 0 | b = Morta 24 0 0 | c = Anormal	a b c <-- classified as 243 1 9 | a = Normal 1 2 0 | b = Morta 22 0 2 | c = Anormal
Avaliação final: 320 amostras		
a b c <-- classified as 277 1 12 | a = Normal 4 0 0 | b = Morta 19 1 6 | c = Anormal	a b c <-- classified as 282 1 7 | a = Normal 3 1 0 | b = Morta 25 1 0 | c = Anormal	a b c <-- classified as 277 1 12 | a = Normal 1 3 0 | b = Morta 24 0 2 | c = Anormal
Avaliação final: 360 amostras		
a b c <-- classified as 319 1 12 | a = Normal 4 1 0 | b = Morta 17 1 5 | c = Anormal	a b c <-- classified as 326 2 4 | a = Normal 3 2 0 | b = Morta 22 1 0 | c = Anormal	a b c <-- classified as 319 3 10 | a = Normal 4 1 0 | b = Morta 21 0 2 | c = Anormal
Avaliação final: 400 amostras		
a b c <-- classified as 353 1 14 | a = Normal 4 1 0 | b = Morta 18 1 8 | c = Anormal	a b c <-- classified as 362 2 4 | a = Normal 3 2 0 | b = Morta 26 1 0 | c = Anormal	a b c <-- classified as 352 4 12 | a = Normal 4 1 0 | b = Morta 22 1 4 | c = Anormal

*Versão A: *Input* de dados espectrais e físicos **Versão B: *Input* de dados espectrais ***Versão C: *Input* de dados físicos

3.4 CONCLUSÕES

O teste ANNSEED proposto mostrou ser rápido e preciso, com uma taxa mínima de acerto de 82,5%.

O ANNSEED mostrou possuir a capacidade de determinar a germinação de sementes de soja, sendo possível em um futuro próximo sua implementação em laboratórios de controle e qualidade de sementes/grãos.

A versão C do ANNSEED mostrou ser a mais promissora dentre todas, pois sua implementação pode ocorrer apenas com um scanner comum e um desktop.

3.5 CONSIDERAÇÕES FINAIS

Existem alguns gargalos inerentes ao processo de desenvolvimento de algoritmos de qualquer espécie, com ANNSEED não é diferente. O mesmo precisa como uma segunda etapa passar por aprimoramento com relação a classificação de sementes anormais e mortas. Isso pode ser solucionado através da expansão de sua biblioteca de dados, uma vez que, até a presente data deste trabalho seu banco de dados possui 400 sementes da cultivar M5947 IPRO.

Outro ponto relevante é o processo de aprendizado do ANNSEED, apensar do mesmo ter passado por alguns testes de exaustão durante a sua confecção, ele precisa ainda passar por um teste com maior amplitude, que só será possível com a expansão de sua biblioteca. Esse teste de exaustão prolongada garantirá que cada nó tenha aprendido o máximo possível, chegando ao seu limite, isso garantirá que as pontes sejam rebalanceadas ao máximo minimizando os erros de classificação em todas as classes.

Outro ponto de melhoria do ANNSEED são suas interfaces, ainda pouco amigáveis. No futuro para sua implementação recomenda-se que o mesmo seja migrado por completo para uma única plataforma integrada de dados. O que tornará sua utilização mais amigável e menos suscetível a erro por parte do usuário.

Agradecimentos

O presente trabalho foi realizado em parte com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Contando também com apoio da FAPESP pela disponibilidade do equipamento multiusuários SeedReporter através do projeto de pesquisa nº 2018/03793-5.

REFERÊNCIAS

ABDAR, M. et al. A review of uncertainty quantification in deep learning: Techniques, applications and challenges. **Information Fusion**, v. 76, p. 243–297, 2021.

ANDRADE, R. G. De et al. Using Computer Vision and Intelligent Classification Techniques for the Classification and Selection of Brazilian Nuts. **Journal of Food Science and Engineering**, v. 6, n. 1, p. 51–62, 2016.

APROSOJA. Acompanhamento da safra de soja. **APROSOJA**, v. 12, p. 1–5, 2019. Disponível em: <http://www.conab.gov.br/OlalaCMS/uploads/arquivos/17_06_08_09_02_48_boletim_graos_junho_2017.pdf>.

BARANAUSKAS, J. A. **Métodos de Amostragem e Avaliação de Algoritmos**. Departamento de Física e Matemática – FFCLRP-USP. Disponível em: <https://dcm.ffclrp.usp.br/~augusto/teaching/ami/AM-I-Metodos-Amostragem.pdf>. Acesso em: 09, janeiro. 2021.

BOUCKAERT, R.R. et al. **WEKA Manual for Version 3-6-8**. University of Waikato, Hamilton, New Zealand, p.303, 2012.

BRASIL. Instrução Normativa n° 45, de 17 de setembro de 2013. **Diário Oficial da República Federativa do Brasil**, 1(183), p.6-27, Brasília, DF, 2013.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. **Regras para análise de sementes.** Brasília: MAPA, 2009.

COELHO, F.; LOVATTO, M.; SILVA, G.B.P. Soja: estresse caro. **Revista Cultivar**, n. 233, p. 22-26, out. 2018.

CONAB - COMPANHIA NACIONAL DE ABASTECIMENTO. Acompanhamento da Safra Brasileira de Grãos: Safra 2020/21. **BRASIL**, v. 8, p. 1–115, 2021.

DE MEDEIROS, A. D. et al. Machine learning for seed quality classification: An advanced approach using merger data from FT-NIR spectroscopy and x-ray imaging. **Sensors (Switzerland)**, v. 20, n. 15, p. 1–12, 2020.

DOBBIN, K. K.; SIMON, R. M. Optimally splitting cases for training and testing high dimensional classifiers. **BMC Medical Genomics**, v. 4, 2011.

DUCOURNAU, S. et al. An image acquisition system for automated monitoring of the germination rate of sunflower seeds. **Computers and Electronics in Agriculture**, v. 44, n. 3, p. 189–202, 2004.

FERREIRA, T.; RASBAND, W. **ImageJ User Guide IJ 1.46r**. NIH - National Institutes of Health. Bethesda, Maryland, 2012.

HONG, P.T.T. et al. Comparative study on vision based rice seed varieties identification. **Seventh International Conference on Knowledge and Systems Engineering**, p. 377-382, 2015. Disponível em:<http://doi.org/10.1109/KSE.2015.46>.

IBM. What is computer vision?. Disponível em: <https://www.ibm.com/topics/computer-vision>. Acesso em: 24, fevereiro. 2021.

LURSTWUT, B.; PORNPANOMCHAI, C. Image analysis based on color , shape and texture for rice seed (Oryza sativa L .) germination evaluation. **Agriculture and Natural Resources**, v. 51, n. 5, p. 383–389, 2017. Disponível em: <https://doi.org/10.1016/j.anres.2017.12.002>.

MARCOS FILHO, J.; KIKUTI, A.L.P.; LIMA, L.B. Métodos para avaliação do vigor de sementes de soja, incluindo a análise computadorizada de imagens. **Revista Brasileira de Sementes**, v.31, n.1, p. 102-112, 2009. Disponível em: <http://dx.doi.org/10.1590/S0101-31222009000100012>.

MARIA, A.; TOMMASELLI, G. Limiarização Automática De Imagens Digitais. **Boletim de Ciências Geodésicas**, v. 6, n. 1, p. 38–48, 2018.

MARQUES FILHO, O.; VIEIRA NETO, H. **Processamento Digital de Imagens**. 1. ed. Rio de Janeiro: Brasport, 1999.

MONTEIRO, R. de C. M. et al. Image processing to identify damage to soybean seeds. **Ciencia Rural**, v. 51, n. 2, p. 1–8, 2020.

OMARY, Z.; MTENZI, F. Machine Learning Approach to Identifying the Dataset Threshold for the Performance Estimators in. **International Journal for Infonomics**, v. 3, n. 3, p. 314–325, 2010.

PAULI, S. T. Z. **DIFFERENTES ABORDAGENS PARA O APRENDIZADO DA REDE NEURAL ARTIFICIAL MULTILAYER PERCEPTRON**. 2020. Universidade do Paraná, 2020. Disponível em: <http://dx.doi.org/10.1016/j.eswa.2016.12.035%0A>.

PEREIRA, G. H. de A.; CENTENO, J. A. S. Avaliação do tamanho de amostras de treinamento para redes neurais artificiais na classificação supervisionada de imagens utilizando dados espectrais e laser scanner. **Boletim de Ciencias Geodesicas**, v. 23, n. 2, p. 268–283, 2017.

RAJU AHMED, M. et al. Classification of pepper seed quality based on internal structure using X-ray CT imaging. **Computers and Electronics in Agriculture**, v. 179, n. October, p. 105839, 2020. Disponível em: <https://doi.org/10.1016/j.compag.2020.105839>.

REIS, L.S. et al. LERCAFÉ: Novo teste para estimar o potencial germinativo de sementes de cafeeiro (*Coffea arabica L.*). **Revista Brasileira de Sementes**, v. 32, n.1, p. 9-16, 2010. Disponível em: <http://dx.doi.org/10.1590/S0101- 31222010000100001>.

RIBEIRO, B. G.; Image analysis of coffee seeds submitted to the LERCAFE test. **Acta Scientiarum Agronomy**, v.38, n.3, 355-361, 2016. Disponível em: <http://dx.doi.org/10.4025/actasciagron.v38i3.28268>.

RIDLER, T. W.; CALVARD, S. Picture Thresholding Using An Interactive Selection Method. **IEEE Transactions on Systems, Man and Cybernetics**, v. smc-8, n. 8, p. 630–632, 1978.

SHARMA, P.; ABROL, P. Color Based Image Segmentation Using Adaptive Thresholding. **International Journal of Scientific and Technical Advancements**, v. 2, n. 3, p. 151–156, 2016. Disponível em: <http://www.ijsta.com/papers/IJSTAV2N2Y16_April/IJSTA-V2N2R53Y16.pdf>.

SCHNEIDER, J. Validação cruzada. Carnegie Mellon University School of Computer Science, 1997. Disponível em: <https://www.cs.cmu.edu/~schneide/tut5/node42.html>. Acesso em: 24, fevereiro. 2021.

SKRUBEJ, U.; ROZMAN, C.; STAJNKO, D. The accuracy of the germination rate of seeds based on image processing and artificial neural networks designed. **Agricultura 12**, v. 1–2, p. 19–24, 2016.

USP. Perceptron Multi-Camadas (MLP). Instituto de Ciências Matemáticas e de Computação – ICMC. Disponível em: <https://sites.icmc.usp.br/andre/research/neural/MLP.htm>. Acesso em: 29, outubro. 2019.

YANG, J. et al. Hyperspectral prediction of sugarbeet seed germination based on gauss kernel SVM. **Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy**, v. 253, n. 74, p. 119585, 2021. Disponível em: <https://doi.org/10.1016/j.saa.2021.119585>.