

LUÍS FELIPE QUEVEDO RAMIREZ

**ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO E ANÁLISE DE IMAGENS
NA AVALIAÇÃO DA QUALIDADE FISIOLÓGICA DE SEMENTES DE ARROZ**

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

Orientador: Denise C. Fernandes dos Santos Dias

**VIÇOSA - MINAS GERAIS
2023**

Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa

T

Q3a
2023

Quevedo Ramirez, Luis Felipe, 1990-

Análise de imagens e espectroscopia no infravermelho próximo na avaliação da qualidade fisiológica de sementes de arroz / Luis Felipe Quevedo Ramirez. – Viçosa, MG, 2023.
1 dissertação eletrônica (55 f.): il. (algumas color.).

Orientador: Denise Cunha Fernandes dos Santos Dias.
Dissertação (mestrado) - Universidade Federal de Viçosa, Departamento de Agronomia, 2023.

Referências bibliográficas: f. 46-55.

DOI: <https://doi.org/10.47328/ufvbbt.2023.337>

Modo de acesso: World Wide Web.

1. *Oriza sativa*. 2. Arroz - Semente - Qualidade - Análise.
3. Espectroscopia de infravermelho próximo. 4. Arroz - Semente - Controle de qualidade - Automação. I. Dias, Denise Cunha Fernandes dos Santos, 1958-. II. Universidade Federal de Viçosa. Departamento de Agronomia. Programa de Pós-Graduação em Agronomia. III. Título.

CDD 22. ed. 633.1821

Bibliotecário(a) responsável: Bruna Silva CRB-6/2552

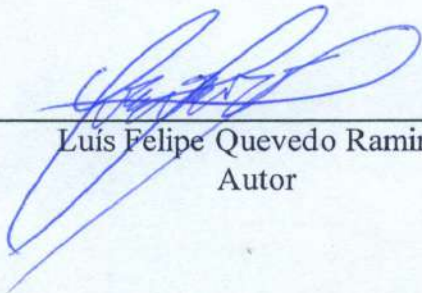
LUÍS FELIPE QUEVEDO RAMIREZ

**ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO E ANÁLISE DE IMAGENS
NA AVALIAÇÃO DA QUALIDADE FISIOLÓGICA DE SEMENTES DE ARROZ**

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

APROVADA: 28 de fevereiro de 2023.

Assentimento:



Luís Felipe Quevedo Ramirez
Autor



Denise Cunha Fernandes dos Santos Dias
Orientadora

A Deus, aos meus pais, ao meu irmão, pelo
incentivo, apoio incondicional e amor.
Dedico.

AGRADECIMENTOS

A Deus, pela sabedoria, inteligência, saúde, amor e estar sempre em todos os caminhos e decisões da minha vida.

Ao meus pais, Ana Paulina e Luís Hernan, por todo o amor, apoio, carinho, esforço e ensinamentos para ser cada dia uma pessoa com grandes valores e confiar sempre em mim para que este sonho seja uma realidade.

Ao meu irmão, Hernan Dario, por todo seu apoio e amor incondicional, por estar sempre a par dos meus projetos e conquistas.

À empresa FEDEARROZ, pela confiança, oportunidade, respaldo e apoio.

À Universidade Federal de Viçosa ao Programa de Pós-graduação em fitotecnia pela minha formação e por todo os valores, conhecimentos, ensinamentos adquiridos e pela oportunidade.

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

À minha professora e orientadora Dra. Denise Cunha F. S. Dias, pelos ensinamentos, compressão, dedicação, por me aconselhar em todas as decisões, pela paciência, apoio e por sempre me guiar no caminho do sucesso.

A meu professor e coorientador Dr. Laercio Júnio da Silva, por sua disponibilidade, observações, ajuda e conhecimentos durante minha formação.

À Julia M. S. por ser aquela pessoa especial e carinhosa, por me apoiar e me guiar em todas as decisões e por fazer parte deste capítulo maravilhoso da minha vida.

Aos meus amigos e técnicos do laboratório de sementes e do Grupo de Estudos em Sementes (GSem): Bruno Noronha, Marcelo, Geovana, Amanda, Joyce, Rafaela, Wander, Joao, Tainara, Daniel, Soryana, Gabriel, Manuel e Jose Custodio pela ajuda e todo o apoio, muito obrigado.

Minha mais sincera gratidão a todos que, de alguma forma direta ou indireta, fizeram possível a realização deste trabalho.

Muito obrigado!

RESUMO

RAMIREZ, Luís Felipe Quevedo, M.Sc., Universidade Federal de Viçosa, fevereiro de 2023. **Espectroscopia no infravermelho próximo e análise de imagens na avaliação da qualidade fisiológica de sementes de arroz.** Orientadora: Denise Cunha Fernandes dos Santos Dias.

Na atualidade, a utilização de tecnologias que permitam avaliar o potencial fisiológico das sementes com precisão, agilidade e sem subjetividade são cada vez mais relevantes. Objetivou-se avaliar a eficiência da análise computadorizada de imagens de plântulas e da espectroscopia no infravermelho próximo para classificar o vigor de lotes de sementes de arroz. Inicialmente, sementes de cinco lotes cv. IRGA 424 IR foram caracterizadas pelos testes de germinação, envelhecimento acelerado, germinação a baixa temperatura e emergência de plântulas. As sementes de cada lote foram submetidas à análise de imagens de plântulas pelo software SAPL®, obtendo-se os parâmetros de comprimento de plântula, de raiz e de parte aérea e os índices de vigor, de uniformidade e de crescimento aos 4, 5 e 7 dias após a semeadura. Pelo software ILASTIK, realizou-se a classificação das plântulas baseada em características morfológicas, com treinamento para o reconhecimento de plântulas normais fortes e fracas e de sementes não germinadas. Duzentas sementes de cada lote foram avaliadas individualmente em um espectrômetro de transformada de Fourier – FT-NIR coletando-se os espectros de refletância NIR (3.112 pontos/espectro) na faixa de comprimento de onda entre 1000 e 2500 nm, com resolução de 8 cm⁻¹. O experimento foi conduzido em DIC com quatro repetições. Foi realizada a análise de variância e as médias obtidas para cada lote comparadas pelo teste de Tukey a 5% de probabilidade. Para os espectros NIR, os espectros foram pré-processados e utilizados para obtenção de modelos de classificação utilizando-se o PSL-DA, sendo 70% dos dados usados para treinamento e 30% para teste. A análise foi feita usando os espectros originais, utilizando diferentes tipos de pré-processamentos que forneceram acurácias relevantes, mas sobressaindo a segunda derivada de Savitzky-Golay. Conclui-se que os softwares SAPL® e ILASTIK foram eficientes na identificação da qualidade fisiológica de sementes de arroz para a avaliação do potencial fisiológico das sementes de arroz com 4 e 5 dias após a semeadura. A espectroscopia do FT-NIR é uma alternativa não destrutiva promissora para a predição rápida do potencial fisiológico das sementes, com acurácia de 96%. As técnicas baseadas em análise de imagem e leituras espectrais fornecem resultados complementares aos demais testes de vigor, demonstrando potencial para classificar os lotes de arroz com rapidez e alta precisão.

Palavras-chave: *Oryza sativa* L. Análise computadorizada. FT-NIR. Automação.

ABSTRACT

RAMIREZ, Luís Felipe Quevedo, M.Sc., Universidade Federal de Viçosa, February, 2023. **Near infrared spectroscopy and image analysis in the evaluation of the physiological quality of rice seeds.** Adviser: Denise Cunha Fernandes dos Santos Dias.

Nowadays, the use of technologies that allow evaluating the physiological potential of seeds with precision, agility and without subjectivity are increasingly relevant. The objective was to evaluate the efficiency of computerized analysis of seedling images and near infrared spectroscopy to classify the vigor of rice seed lots. Initially, seeds from five lots cv. IRGA 424 IR were characterized by germination, accelerated aging, cool germination and seedling emergence tests. The seeds of each lot were submitted to the analysis of seedling images by the SAPL® software, obtaining the parameters of seedling length, root and aerial part and the indices of vigor, uniformity and growth at 4, 5 and 7 days after sowing. The images were classified using the ILASTIK software, based on a color scale and pixels, with training for the recognition of strong and weak normal seedlings and dead seeds. Two hundred seeds of each lot were evaluated individually in a Fourier transform spectrometer – FT-NIR, collecting NIR reflectance spectra (3,112 points/spectrum) in the wavelength range between 1000 and 2500 nm, with a resolution of 8 cm⁻¹. The experiment was conducted in DIC with four replications. Analysis of variance was performed and the means obtained for each lot compared by Tukey's test at 5% of probability. For the NIR spectra, PLS-DA techniques were employed, with 70% of the data used for training and 30% for testing. The analysis was performed using the original spectra, compared to the spectra processed through the second derivative of the Savitzky-Golay filter. It was concluded that the SAPL® and ILASTIK software were efficient for evaluating the physiological potential of rice seeds four and five days after sowing. FT NIR spectroscopy is a promising non-destructive alternative for the rapid prediction of the physiological potential of seeds, with an accuracy of 96%. Techniques based on image analysis and spectral readings showed results similar to the other vigor tests, demonstrating the potential to classify rice lots quickly and with high precision.

Keywords: *Oryza sativa* L. Computer analysis. FT-NIR. Automation

LISTA DE ILUSTRAÇÕES

Figura 1. Aquisição das imagens fotográficas das plântulas e processamento por meio do software SAPL®.....	21
Figura 2. Etapas do processamento das imagens de sementes de arroz pelo software ILASTIK. A) plântulas normais fortes; B) plântulas normais fracas; C) plântulas anormais; D) sementes não germinadas.	23
Figura 3. Probabilidade de os lotes de sementes gerarem plântulas normais fortes (PNF), plântulas normais fracas (PNFra), plântulas anormais (PA) e sementes não germinadas (SNG) no 4° (A), 5° (B) e 7° (C) dia após a semeadura, de acordo com dados obtidos pelo software ILASTIK.....	34
Figura 4. Germinação e classificação da qualidade fisiológica de cinco lotes de arroz cv. IRGA 424.	35
Figura 5. Espectros FT-NIR de diferentes lotes de sementes de arroz cv. IRGA 424: Espectros originais (A) espectros obtidos para cada classes de qualidade de semente (B) espectros médios originais.	36
Figura 6. Análise de componentes principais usando os dados originais FT-NIR obtidos de sementes de diferentes lotes de arroz cv. IRGA 424. Diagrama de ordenação central (A) e círculo de correlação e distribuição (B).	37
Figura 7. Gráfico de pontuação da análise discriminante por mínimos quadrados parciais (PLS-DA) com os dois primeiros componentes discriminantes de sementes de arroz com base nos espectros FT-NIR transformados pela segunda derivada Savitzky-Golay. As elipses mostram intervalos de confiança de 95% em cada classe de sementes.	40
Figura 8. Matriz de confusão da classificação dos lotes de sementes de arroz cv. IRGA 424 quanto ao potencial fisiológico após o pré-processamento com a segunda derivada de Savitzky-Golay.....	41
Figura 9. Importância das variáveis de comprimento de ondas para a classificação via PLS-DA dos níveis de vigor dos lotes de arroz cv. IRGA 424.	42

LISTA DE TABELAS

Tabela 1 – Caracterização da qualidade fisiológica de cinco lotes de sementes de arroz cv. IRGA 424 avaliados pelo grau de umidade (GU), germinação (G), primeira contagem de germinação (PCG), envelhecimento acelerado (EA), índice de velocidade de emergência (IVE) e germinação a baixa temperatura (GBT).	26
Tabela 2. Resultados médios de comprimento de parte aérea (CPA), comprimento radicular (CR), comprimento total da plântula (CTP), índice de uniformidade (IU), índice de crescimento (IC) e índice de vigor (IV), obtidos pelo SAPL® aos quatro, cinco e sete dias após a semeadura de cinco lotes de sementes de arroz cv. IRGA 424.	30
Tabela 3. Resultados de acurácia e confiabilidade (kappa) para o conjunto de dados de treinamento e validação (teste) dos modelos de classificação gerados por meio da análise PLS-DA com diferentes tipos de pré-processamentos em espectros FT-NIR de sementes de cinco lotes de arroz cv. IRGA 424.	38

SUMÁRIO

1. INTRODUÇÃO.....	11
2. REVISÃO DE LITERATURA.....	12
2.1. Análise computadorizada de plântulas para avaliação do vigor em sementes	12
2.2. Espectroscopia do Infravermelho Próximo (NIR)	16
3. MATERIAIS E MÉTODOS.....	18
3.1. Caracterização da qualidade inicial dos lotes	18
3.1.1. Determinação do grau de umidade	18
3.1.2. Determinação do grau de umidade	18
3.1.3. Determinação do grau de umidade	19
3.1.4. Teste de germinação a baixa temperatura.....	19
3.1.5. Envelhecimento acelerado	19
3.1.6. Emergência de plântulas	19
3.1.7. Índice de velocidade de emergência (IVE).....	19
3.1.8. Delineamento experimental e análise estatística	19
3.2. Ensaio I – Análise computadorizada de plântulas pelos softwares SAPL® e ILASTIK para avaliação do potencial fisiológico das sementes	20
3.2.1. Análise computadorizada de plântulas	20
3.2.2. Software SAPL®.....	21
3.2.3. Software ILASTIK.....	22
3.2.4. Delineamento experimental e análise estatística	23
3.3. Ensaio II – Espectroscopia no infravermelho próximo (NIR) e sua relação com o potencial fisiológico das sementes	24
3.3.1. Aquisição dos espectros FT-NIR	24
3.3.2 Delineamento experimental e análise estatística	24
4. RESULTADOS E DISCUSSÃO.....	25
4.1. Caracterização da qualidade inicial dos lotes de arroz	25
4.1. Ensaio I – Análise computadorizada de plântulas pelos softwares SAPL® e ILASTIK para avaliação do potencial	27
4.2.1. Análise de plântulas pelo software SAPL®.....	27
4.2.2. Análise de plântulas pelo software ILASTIK.....	31
4.2. Ensaio II – Espectroscopia no infravermelho próximo (NIR) e sua relação com o potencial fisiológico das sementes	35
5. CONCLUSÕES	45
REFERÊNCIAS	46

1. INTRODUÇÃO

O arroz (*Oryza sativa* L.) é o segundo cereal mais cultivado no mundo, com aproximadamente 161 milhões de hectares, e produção de cerca de 765,5 milhões de toneladas de grãos em casca, o que corresponde a 29% do total de grãos usados na alimentação da população mundial (SOSBAI, 2018). No Brasil, o maior estado produtor é o Rio Grande do Sul, representando cerca de 73% de toda a produção do país e ocupando em torno de 66% da área de cultivo nacional (CONAB, 2022). A estimativa de produção para o ano comercial de 2022/2023 para o Brasil é de 10,4 milhões de toneladas e área semeada de 1,5 milhões de hectares, prevendo uma redução da área de 9,3% em comparação com a safra anterior (CONAB, 2023).

O arroz é uma gramínea anual, de mecanismo fotossintético C3 e adaptada ao cultivo irrigado, de várzea e sequeiro. A variabilidade climática pode interferir no crescimento, desenvolvimento e na produtividade da cultura. Em condições desfavoráveis, a formação da semente pode ser prejudicada e afetar seu potencial germinativo. A utilização de sementes de alta qualidade fisiológica pode garantir estande mais uniforme e, consequentemente, maior produtividade (ZONTA et al., 2022).

A qualidade fisiológica das sementes é representada pela germinação, vigor e longevidade. A germinação é avaliada sob condições ideais, padronizadas em laboratório, de modo a obter a porcentagem máxima de plântulas normais, conforme instruções das Regras para Análise de Sementes (BRASIL, 2009). Torna-se necessário, portanto, a avaliação do vigor já que nem sempre os resultados de germinação se correlacionam com a emergência de plântulas no campo e com o desempenho das sementes durante o armazenamento (MARCOS-FILHO, 2015). Assim, para se obter informações mais detalhadas e seguras sobre o desempenho dos lotes e para um controle de qualidade mais efetivo quanto ao destino destes antes da comercialização, é importante que sejam conduzidos testes de vigor, os quais complementam as informações fornecidas pelo teste de germinação.

Segundo Marcos-Filho (2010) é importante a utilização de técnicas inovadoras que permitam a padronização e a redução do tempo para a condução dos testes de vigor. Na atualidade, a avaliação do potencial fisiológico das sementes deve ser feita preferencialmente por meio de testes rápidos que possibilitem a otimização de recursos computacionais, fornecendo resultados precisos e confiáveis em menor tempo (CASTAN; GOMES-JUNIOR; MARCOS-FILHO, 2018). Dentre estes testes, destaca-se o teste de crescimento de plântulas

que tradicionalmente é feito de forma manual e morosa, podendo estar sujeito a erros de interpretação (MARCOS- FILHO, 2015).

Neste contexto, a análise computadorizada de imagens de plântulas é uma alternativa que favorece a condução do teste em menor tempo em comparação com outros procedimentos, já que tem a possibilidade de determinar características difíceis de serem obtidas de forma manual (ALVARENGA et al., 2012).

Segundo Medeiros; Pereira; Silva (2018) há softwares modernos, de fácil acesso, alguns gratuitos e outros que podem ser adquiridos a baixo custo, os quais permitem a avaliação da qualidade da semente e que tem alto desempenho na análise computadorizada. Os softwares Sistema de Análise de Plântulas (SAPL®) e o ILASTIK (MEDEIROS e PEREIRA, 2018; CUA- DRADRO, 2020; MEDEIROS et al., 2020a) são programas computacionais de livre acesso para o processamento de imagens as quais são obtidas por meio de uma captura digital (LIMÃO, 2022).

Outra técnica que pode ser uma alternativa interessante para a avaliação do potencial fisiológico das sementes é a espectroscopia do infravermelho próximo (NIR). Em síntese, esta técnica baseia-se na absorção em comprimentos de onda e de radiação eletromagnética nas faixas entre 780-2500 nm, absorvidos pela água e por compostos químicos como carboidratos, lipídios e proteínas presentes nas sementes, possibilitando assim a formação de espectros correspondentes à absorção de energia entre os grupos funcionais (XIA et al., 2019). É um método rápido, não destrutivo, cujos resultados relacionam as alterações na composição química das sementes ao seu potencial fisiológico (XIA et al., 2019; RIBEIRO et al., 2020; MEDEIROS et al., 2020b).

Diante do exposto, com este trabalho objetivou-se: i) avaliar procedimentos adequados para a identificação de diferenças no potencial fisiológico de lotes de sementes de arroz pelos softwares SAPL® e ILASTIK; ii) determinar se o método da espectroscopia no infravermelho próximo FT-NIR tem potencial para a caracterização do potencial fisiológico das sementes.

2. REVISÃO DE LITERATURA

2.1. Análise computadorizada de plântulas para avaliação do vigor em sementes

A utilização de métodos computadorizados para obter informações e resultados baseados em imagens tem sido identificado como uma abordagem promissora e capaz de capturar e quantificar informações sobre a qualidade fisiológica da semente (MEDEIROS; PEREIRA, 2018). Na atualidade, as tecnologias e ferramentas inovadoras para a avaliação do vigor das sementes se intensificam e são indispensáveis para que as empresas produtoras de sementes possam obter informações confiáveis quanto à qualidade dos lotes destinados à comercialização, seja quanto ao seu desempenho em campo ou no armazenamento (ARAÚJO et al., 2021).

Nos laboratórios de sementes, a automatização dos procedimentos relacionados à análise de sementes que possibilitem capturar e quantificar informações referentes ao seu potencial fisiológico é importante para reduzir e de possíveis erros cometidos por avaliações subjetivas feitas por analistas, os quais avaliam grandes quantidades de lotes de sementes (MEDEIROS; PEREIRA, 2018). Estas tecnologias permitem aumentar a agilidade do processo de avaliação de sementes, proporcionando alta confiabilidade nas informações obtidas (MARCOS-FILHO, 2015).

A análise de imagens de plântulas permite avaliar o vigor das sementes com base em características morfológicas e cálculo de índices de crescimento de plântulas, como vantagens, fornece resultados rápidos e confiáveis na avaliação de plântulas e permitem obter um banco de dados e informações que podem ser armazenados e consultados quando necessário (GOMES-JUNIOR, 2020; MARCOS-FILHO, 2015a). McCormac et al. (1990) realizaram os primeiros estudos de análise computadorizada de plântulas por meio das avaliações de comprimento de raiz primária de tomate, alface e couve-flor. A eficiência das análises computadorizadas já foi investigada para a avaliação do vigor de sementes com sucesso em espécies como abóbora (Silva et al. 2017), milho (FERREIRA et al., 2020; CASTAN et al., 2018), trigo (BRUNES et al., 2016), girassol (Rocha et al., 2015), amendoim (BARBOSA et al., 2016), feijão (ABUD et al., 2022), grão-de-bico (ARAÚJO et al., 2021a), soja (ANTUNES-NETO et al., 2020), arroz (BRUNES et al., 2019; CUADRADO, 2020), entre outros.

A tecnologia de análise de imagem e plântulas tem permitido a adaptação e desenvolvimento de métodos inovadores, rápidos, de acurácia alta e, na maioria das vezes não destrutivos, a partir da criação de modelos de aprendizado de máquina baseados em atributos fisiológicos, sanitários e físicos dos lotes de sementes, que podem ser analisados em tempo real de forma objetiva (GOMES-JUNIOR, 2020). Além disso, a utilização destas ferramentas contribui para a redução do período necessário para a condução do teste.

Diversos softwares foram desenvolvidos para auxiliar na análise computadorizada de sementes e plântulas permitindo fazer inferências sobre o vigor dos lotes. Dentre eles, destacam-se o Seed Vigor System (SVIS®), o Sistema Automatizado do Vigor de Sementes (Vigor-S®), o GroundEye®, o ImageJ, o Sistema de Análise de Plântulas (SAPL®) e o ILASTIK (SAKO et al., 2001; CASTAN et al., 2018; TBIT, 2015; COLLINS, 2007; MEDEIROS; PEREIRA, 2018; BERG et al., 2019).

A Universidade Federal do Rio Grande do Norte desenvolveu o software SAPL®, o qual pode ser adquirido gratuitamente para a avaliação do vigor de sementes e as imagens podem ser obtidas por câmera digital ou dispositivo móvel, como smartphones. A partir da captura das imagens, o software permite obter parâmetros como o comprimento total de plântulas, comprimento radicular e da parte aérea além de gerar índices de vigor de uniformidade e de crescimento de plântulas (MEDEIROS; PEREIRA, 2018).

O índice de vigor é baseado na velocidade e uniformidade de desenvolvimento das plântulas em relação ao máximo valor estimado com a idade pré-estabelecida. A uniformidade é definida com base nos desvios de comprimento de cada plântula, em relação ao tamanho máximo desenvolvido estimado. Assim, após o processamento das imagens, o software gera automaticamente valores numéricos, ou seja, índices referentes ao vigor, crescimento e à uniformidade das plântulas (REIS, 2017).

Estudos com diferentes espécies como arroz (CUADRADO, 2020), moringa (PEREIRA et al., 2020), milho (MEDEIROS et al., 2019), grão-de-bico (ARAÚJO et al., 2021a), lentilha (LIMAO, 2022) e soja (MEDEIROS e PEREIRA, 2018) confirmam a eficiência do software SAPL® para a avaliação do potencial fisiológico de sementes. Em arroz, CUADRADO (2020), ao avaliar plântulas pelo SAPL com quatro e cinco dias após o início do teste de germinação, observou o SAPL não foi tão eficiente aos quatro dias, principalmente para destacar o lote de pior desempenho, obtendo resultados comparáveis a outros testes de vigor quando as avaliações das plântulas foram feitas aos cinco dias. A análise de imagens de plântulas pode ser realizada utilizando-se o software ILASTIK, que envolve uma abordagem de aprendizado de máquina orientada ao usuário é o Interactive Machine Learning (IML). As abordagens IML são definidas como modelos interativos de aprendizagem automático que permitem ao usuário classificar e analisar imagens (BERG, 2019). Assim, o ILASTIK destaca-se nesta categoria, por ser uma ferramenta para a classificação que utiliza as abordagens IML para a segmentação e análise de imagens interativas.

O ILASTIK é um software de fácil execução e manipulação, permite a segmentação e classificação de dados de acordo com rótulos fornecidos pelo usuário, sendo um software de

processamento de imagens de autoaprendizado (SOMMER et al., 2011). O software é ideal para usuários que não têm um conhecimento amplo de computação (BERG et al., 2019).

Na área de sementes, este software tem sido utilizado para a avaliação de plântulas obtidas no teste de germinação que podem ser classificadas quanto ao vigor conforme critérios estabelecidos pelo analista.

Assim, o ILASTIK apresenta uma interface dinâmica, permitindo que o usuário estabeleça de forma livre e arbitrária a segmentação das classes de plântulas como por exemplo, plântulas fortes e fracas. Neste contexto, o usuário tem a autonomia de treinar classificadores aleatórios baseados nos recursos de imagens, de acordo com o objetivo da sua avaliação. O software é acoplado em módulo CellProfiler (gratuito e de acesso livre), que possibilita a classificação de bioimagens de forma interativa (LIMÃO, 2022). No pré-processamento e na segmentação das imagens é utilizada a ferramenta de classificação de pixels, definindo duas classes de segmentação “semente ou plântula” e “fundo”, baseadas na intensidade de cores e pixels. Para criar os mapas de probabilidade, os pixels pertencentes às regiões de cada classe de segmentação são treinados diretamente nos dados de entrada para a identificação das plântulas conforme as cores das classes de segmentação. Assim, o software é treinado para o reconhecimento de classes pré-estabelecidas de plântulas (MEDEIROS et al., 2020; CUADRADO, 2020).

Este software foi aplicado, por exemplo, na triagem de alto rendimento para quantificar danos causados por tripes (Thysanoptera: Thripidae) em folhas de pimenta (Capsicum) (VISSCHERS et al., 2018), na quantificação e análise espacial de características em imagens histológicas de cérebro de roedores (YATES et al., 2019) e na classificação de vigor de sementes de soja (MEDEIROS et al., 2020a) e arroz (CUADRADO, 2020). Segundo Medeiros et al. (2020), o software ILASTIK, teve alta acurácia para classificar sementes de soja quanto à presença de danos e também classificar as plântulas quanto ao nível de vigor. Os resultados obtidos em plântulas de arroz permitiram a estratificação dos lotes em níveis de qualidade fisiológica fornecendo informações coerentes à classificação obtida nos testes de vigor tradicionais utilizados.

Em síntese, na atualidade, a utilização de softwares de livre acesso e uso de inteligência artificial para a análise de imagem de plântulas é fundamental para o avanço do conhecimento em ciência e tecnologia de sementes. Estas tecnologias conferem eficiência e precisão nas avaliações garantindo confiabilidade no processo de tomada de decisões para o controle de qualidade das sementes (MEDEIROS et al., 2020b).

2.2. Espectroscopia do Infravermelho Próximo (NIR)

A espectroscopia no infravermelho próximo (NIR) é uma técnica utilizada para quantificar e identificar informações químicas, sendo um método não destrutivo e rápido (RAVAZZI, 2019). Esta técnica já é empregada à produtos agrícolas, alimentos, nas indústrias farmacêuticas, petroquímicas e químicas (BLANCO E VILLARROYA, 2002).

No campo da tecnologia de sementes buscam-se metodologias inovadoras, rápidas, não destrutivas e sem subjetividade que permitam obtenção de resultados correlacionados com os de outros testes indicadas para a avaliação da qualidade fisiológica das sementes. Neste contexto, a espectroscopia no infravermelho próximo (NIR) é uma ferramenta com potencial para classificação das sementes em níveis de qualidade fisiológica baseada na detecção de seus componentes químicos (REDDY, 2022).

A espectroscopia no infravermelho próximo (NIR) é baseada na absorção de radiação eletromagnética em comprimentos de onda na faixa de 780-2500 nm. Essa faixa fornece informações relacionadas à composição do material de estudo, de modo que a leitura dos espectros permite identificar presença de grupos funcionais (C-H, N-H, O-H). Essa leitura depende da interação do composto com a radiação eletromagnética, sendo absorvida por água, carboidratos, lipídios e proteínas (LARIOS, 2020).

Assim, levando-se em consideração que a composição química das sementes pode ser alterada devido ao processo de deterioração, a partir dos espectros obtidos, pode-se fazer uma estimativa do potencial fisiológico das sementes pela técnica NIR (LIMAO, 2022).

Porém, segundo Medeiros (2021), nem sempre esses espectros originais obtidos podem se correlacionar diretamente com a sua composição química e qualidade fisiológica, sendo necessária aplicar técnicas adequadas para o processamento dos dados bem como escolher métodos assertivos de classificação.

Assim, os dados originais obtidos pelo equipamento podem necessitar da aplicação de tratamentos de pré-processamento para reduzir a variabilidade, remover ruídos ou fenômenos físicos dos espectros a fim de melhorar o modelo de regressão, de classificação ou de análise exploratória subsequente. Segundo Pasquini (2018), a aplicação dos algoritmos de pré-processamento dos espectros obtidos deve ser feita com precaução, considerando a característica do espectro, já que se pode ter perda de informações importantes quando o método de pré-processamento for inadequado.

Dentro das técnicas de pré-processamento mais comumente utilizadas para tratamento dos espectros destacam-se dois grupos: os métodos de correção de dispersão e métodos deriva-

tivos. Dentro dos métodos de correção de dispersão as mais comuns são a correção de espalhamento multiplicativa (Multiplicative Scatter Correction - MSC), a variável normal padrão (Standard Normal Variate - SNV) e a normalização, utilizadas para a diminuição dos efeitos aditivos e multiplicativos gerados pelo espalhamento, pois os algoritmos corrigem o deslocamento e tendência da linha de base (NOLASCO, 2019).

Dentre os métodos derivativos, destaca-se a derivada de Savitzky-Golay, cujo objetivo principal remover efeitos aditivos e multiplicativos nos espectros, em que a primeira derivada remove a linha de base, enquanto a segunda derivada remove a linha de base e a tendência linear. Portanto, permite suavizar os dados aumentando a precisão sem alterar a tendência do sinal, corrigindo ao mesmo tempo o problema de ruído, deslocamento vertical e inclinação de linha de base (RINNAN et al., 2009).

Para o desenvolvimento dos modelos de classificação dos lotes de sementes quanto ao vigor, há necessidade de reconhecimento de padrão ou agrupamento desenvolvidos com base nos espectros NIR da amostra (MORTENSEN, 2021).

Burns e Donald (2007) relatam que para o controle de qualidade em sementes, é necessária a combinação dos diferentes métodos matemáticos e estatísticos, como análises multivariadas, análises de componentes principais (PCA) e a regressão parcial de mínimos quadrados (PLS), para ter uma correlação entre os dados espectrais e a composição química das sementes. Segundo Kumar et al. (2014), a implementação dos dados multivariados tem a probabilidade de diminuir a quantidade de ruídos aumentando a precisão aos métodos e possibilitando a determinação contínua de diversas características da amostra com um único sinal multivariado, permitindo, assim, modelar a relação entre os dados espectrais e a composição química de determinada amostra ou separação em classes.

Uma vez concluída a etapa de modelagem, é pertinente validar o modelo, para observar se foi capaz ou não de analisar corretamente os parâmetros desejados, partindo das validações internas e externas. Para a validação interna de um conjunto de determinadas amostras, se aparta um conjunto delas as quais não entram nesta calibração, e é feita uma previsão das características dessa amostra. Para a avaliação externa, se processam aquelas amostras que não tiveram participação da construção da validação interna (BERRUETA et al., 2007).

O NIR é um método com o qual se pode obter um grande número de informações em um período curto de tempo, reduzindo a mão de obra, não poluente e não precisa a utilização de produtos químicos ou reagentes (AMORIM, 1996; MAYRINCK, 2018).

Na área de tecnologia de sementes há relatos de que o NIR permite diferenciar sementes vazias ou malformadas daquelas com preenchimento adequado, auxiliando na avaliação da qua-

lidade física dos lotes (FARHADI et al., 2015) e também na avaliação da pureza física de sementes de arroz (ZHANG et al., 2019). Também vem sendo testado para avaliar a qualidade fisiológica e viabilidade de sementes de diversas espécies como grão-de bico (RIBEIRO et al., 2020), lentilha (LIMÃO, 2022), soja (WANG et al., 2017; BAZONI et al., 2017), trigo (XIA et al., 2019; ZHU et al., 2019), milho (QIU et al., 2018), melão (KANDPAL et al., 2016), algodão (GAITÁN et al., 2008; HUANG et al., 2013), café (GUIMARÃES, 2016) e tomate (SHRES-THA et al., 2016).

Além disto, há estudos avaliando a eficiência da utilização do NIR na identificação de misturas genéticas em lotes de sementes, assim como para a identificação de sementes de outras espécies (SHRESTHA et al., 2016; VREŠAK et al., 2016 e NIE et al., 2019).

3. MATERIAIS E MÉTODOS

A pesquisa foi conduzida no Departamento de Agronomia da Universidade Federal de Viçosa, no Laboratório de Pesquisa de Sementes, Viçosa, Minas Gerais, Brasil. Foram utilizados cinco lotes de sementes de arroz da cultivar IRGA 424 RI, colhidos na safra de 2020. Inicialmente, as sementes de cada lote foram acondicionadas em sacos de papel kraft e mantidas em câmara fria a 10 ± 2 °C e a 50% UR durante todo o período experimental. Cada lote de sementes foi submetido aos seguintes testes para a caracterização inicial da sua qualidade.

3.1. Caracterização da qualidade inicial dos lotes

3.1.1. Determinação do grau de umidade

Foi realizado pelo método da estufa a 105°C durante 24 horas (BRASIL, 2009), com duas repetições de 50 sementes para cada lote. Os resultados foram expressos em porcentagem média com base no peso úmido (b.u.).

3.1.2. Determinação do grau de umidade

Realizado conforme as Regras para Análise de Sementes (BRASIL, 2009) com quatro repetições de 50 sementes em rolos de papel toalha Germitest® umedecido com quantidade de água equivalente a 2,5 vezes o peso do papel seco. Os rolos foram mantidos em germinador de sementes à temperatura de 25°C. As avaliações foram realizadas aos 5 e 14 dias após a semeadura e os resultados expressos em porcentagem média de plântulas normais.

3.1.3. Determinação do grau de umidade

Foi efetuado juntamente com o teste de germinação, calculando-se a porcentagem de plântulas normais obtidas no quinto dia após a semeadura (BRASIL, 2009).

3.1.4. Teste de germinação a baixa temperatura

Foi realizado conforme descrito por Dias e Marcos-Filho (2020) com quatro repetições de 50 sementes em rolos de papel toalha Germitest® umedecido com a quantidade de água equivalente a 2,5 vezes o peso do papel seco. Os rolos foram mantidos em germinador de sementes à temperatura de 18°C. As avaliações foram realizadas aos 9 dias após a semeadura, e os resultados foram expressos em porcentagem de plântulas com tamanho maior ou igual a 5 cm.

3.1.5. Envelhecimento acelerado

Conduzido com quatro repetições de 50 sementes para cada tratamento distribuídas sobre uma tela metálica acoplada à caixa plástica tipo gerbox, contendo 40 ml de água destilada e mantidas em incubadora BOD a 42 °C durante 120 h (MARCOS-FILHO et al., 2020, 1999). As caixas foram tampadas para obter aproximadamente 100% de UR no seu interior. Após esse período, as sementes foram colocadas para germinar conforme da maneira descrita acima para o teste de germinação e a avaliação realizada no quinto dia após a semeadura. Os resultados foram expressos em porcentagem de plântulas normais.

3.1.6. Emergência de plântulas

O teste foi realizado em sala de crescimento vegetal em bandejas plásticas contendo mistura de solo e areia na proporção de 2:1, com quatro repetições de 50 sementes semeadas na profundidade de 1,0 cm. A porcentagem de emergência foi calculada computando-se o total de plântulas emergidas, com o tamanho igual ou superior de 3 cm acima da superfície do substrato, após a completa estabilização do estande.

3.1.7. Índice de velocidade de emergência (IVE)

Conduzido juntamente com o teste de emergência de plântulas descrito acima, realizando-se contagens diárias do número de plantas emergidas, do primeiro ao último dia de duração do teste. O IVE foi calculado conforme Maguire (1962).

3.1.8. Delineamento experimental e análise estatística

Os testes foram conduzidos em delineamento inteiramente casualizado (DIC), com cinco lotes e quatro repetições para todos os testes. Os dados foram testados quanto à

distribuição normal dos erros pelo teste de Shapiro-Wilk e à homogeneidade de variâncias pelo teste de Bartlett. Os dados referentes para a caracterização inicial da qualidade fisiológica dos lotes e análise computadorizada de plântulas foram submetidos à análise de variância para significância do teste F, e as médias obtidas para cada lote foram comparadas pelo teste de Tukey a 5% de probabilidade. Para todas as análises, foi utilizado o software estatístico R 4.1.1. (R CORE TEAM, 2020).

3.2. Ensaio I – Análise computadorizada de plântulas pelos softwares SAPL® e ILASTIK para avaliação do potencial fisiológico das sementes

3.2.1. Análise computadorizada de plântulas

Inicialmente, para aquisição das imagens foi feito o teste de comprimento de plântulas com quatro repetições de 20 sementes por lote, distribuídas equidistantemente sobre uma linha traçada no terço superior do papel toalha, no sentido longitudinal, segundo Nakagawa *et al.* (1999). O teste foi conduzido em germinador a 25 o C. Foram avaliadas as plântulas com quatro, cinco e sete dias. Ao final de cada período, as plântulas foram transferidas do papel de germinação para uma folha de espuma vinílica acetinada (E.V.A) de coloração azul (40 x 60 cm), contendo onze células de cinco centímetros de largura, divididas por faixas brancas e, em seguida, fotografadas (Figura 1). A aquisição das imagens foi feita com uma câmera digital Canon, modelo EOS Rebel T7, de 24.1 Megapixels, com 1/60 segundos de velocidade de disparo do obturador e F4.0 de abertura do diafragma. A câmera foi mantida a 45 cm das amostras e angulação de 90° em relação a base fotográfica, utilizando suporte do tipo *Copystand*.



Figura 1. Aquisição das imagens fotográficas das plântulas e processamento por meio do software SAPL®.

3.2.2. Software SAPL®

As imagens das plântulas foram armazenadas e, posteriormente, inseridas no arquivo default selecionado do software SAPL®. O software forneceu medições do comprimento da parte aérea - CPA (mm plântula-1), comprimento radicular – CR (mm plântula-1) e comprimento total da plântula – CTP (mm plântula-1). Posteriormente, os valores de comprimento de plântulas obtidos pelo SAPL® foram inseridos no pacote SeedCalc do software R (SILVA et al 2019) para gerar índices de uniformidade (CASTAN et al., 2018), crescimento (MEDEIROS et al., 2019) e vigor (MEDEIROS; PEREIRA, 2018), conforme as fórmulas descritas abaixo:

$$IC = \frac{(médiaCT)}{X_{max}} \times 1000$$

Em que $IC = (\text{médiaCT}/X_{\text{máx}}) \times 100$; IC= índice de crescimento; media CT = media do comprimento total de plântula da repetição; e $X_{\text{máx}}$ = media do comprimento total de 5% das maiores plântulas do experimento.

$$\text{Uniformity_2} = \left[1 - \frac{\sum_{i=1}^n |X_i - \bar{X}|}{n \times X} \right] \times 1000 - \left[n_{\text{dead}} \times \left(\frac{50}{n_{\text{total}}} \right) \right]$$

Em que: *Uniformity_2* é o índice de uniformidade; X_i , o comprimento da plântula analisada; X , o comprimento médio de plântulas do lote de sementes analisado; n , a variável número de plântulas totais avaliadas; n_{dead} , o número de sementes não germinadas ou mortas presentes no lote avaliado; e n_{total} , o número total de plântulas.

$$\text{Vigor_Corr} = [(Growth \times wg) + (Uniformity \times wu)] \times \left(\frac{G}{100} \right)$$

Em que: *Vigor_Corr* é o índice de vigor corrigido; Growth é o índice de crescimento e Uniformity é o índice de uniformidade escolhido pelo usuário; wg e wu são pesos ajustáveis na fórmula para o crescimento e uniformidade, porém com valores referência de 70 e 30, respectivamente; G é o percentual de germinação do lote.

3.2.3. Software ILASTIK

Neste software, se tem a autonomia de treinar classificadores aleatórios baseados nos recursos de imagens de acordo com o objetivo proposto. O software é acoplado em módulo *CellProfiler*, o qual é gratuito e de acesso livre, que possibilita a classificação de bioimagens de forma interativa. As imagens obtidas conforme descrito acima para o software SAPL® foram carregadas no software ILASTIK. Para o pré-processamento e a segmentação das imagens foi utilizada a ferramenta de classificação de pixels, definindo duas classes de segmentação (semente ou plântula e fundo), baseadas nas características de coloração. Seguidamente o software gera as imagens em escala de cinza e realizar é realizado então o treinamento dos dados de entrada para a identificação das plântulas conforme as cores das classes de segmentação. O software foi treinado para o reconhecimento de quatro classes: as plântulas normais fortes (PNF), que são aquelas plântulas morfologicamente saudáveis mostrando crescimento vigoroso da parte aérea e parte radicular; plântulas normais fracas (PNFra), as quais se mostraram subdesenvolvidas em relação as plantas vigorosas e fortes; plântulas anormais (PA), aquelas que não mostram potencial para continuar seu desenvolvimento com qualquer estrutura essencial ausente ou muito danificada, com distúrbios fisiológicos, estruturas

desproporcionais, com infecção primária e sementes não germinadas (SNM). O classificador treinado foi aplicado a todas as imagens e os mapas de probabilidade foram exportados para cada imagem, gerando dados de número de plântulas normais fortes (PNF), número de plântulas normais fracas (PNFra), número de plântulas anormais (PA) e sementes não germinadas (SNM) (Figura 2).

As etapas de processamento do método interativo aplicado à classificação de plântulas de arroz são apresentadas na Figura 2. A aquisição das imagens foi feita utilizando-se um fundo azul, onde as plântulas foram individualmente dispostas em cada célula para facilitar a separação das mesmas. Em seguida, foi aplicada a segmentação da identificação das sementes e plântulas para melhorar o mapa de probabilidade, sendo que o software identifica através do mapa de probabilidade sementes e plântulas individuais. A previsão da classificação para cada semente e plântula foi diferenciada por cada cor de treinamento segundo seu grupo de classificadores.

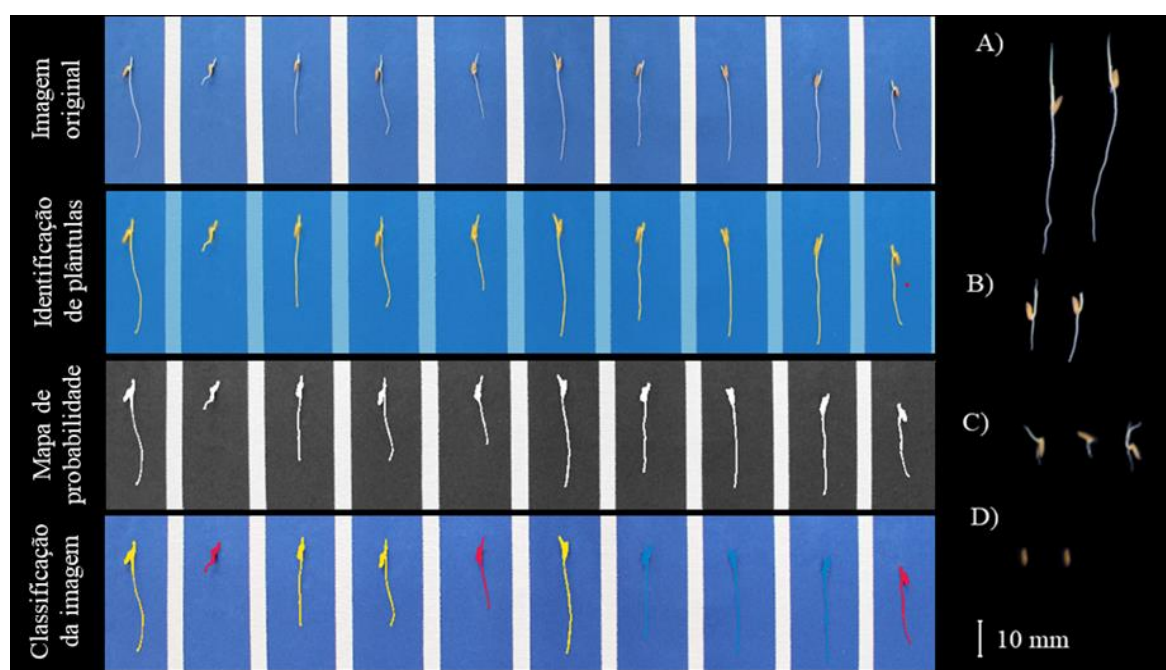


Figura 2. Etapas do processamento das imagens de sementes de arroz pelo software ILASTIK. A) plântulas normais fortes; B) plântulas normais fracas; C) plântulas anormais; D) sementes não germinadas.

3.2.4. Delineamento experimental e análise estatística

Adotou-se o delineamento inteiramente casualizado (DIC), com cinco lotes e quatro repetições. Os dados foram testados quanto à distribuição normal dos erros pelo teste de Shapiro-Wilk e à homogeneidade de variâncias pelo teste de Bartlett. Em seguida, os dados

foram submetidos à análise de variância para significância do teste F, e as médias obtidas para cada lote foram comparadas pelo teste de Tukey a 5% de probabilidade. Para todas as análises, foi utilizado o software estatístico R 4.1.1. (R CORE TEAM, 2020).

3.3. Ensaio II – Espectroscopia no infravermelho próximo (NIR) e sua relação com o potencial fisiológico das sementes

3.3.1. Aquisição dos espectros FT-NIR

Utilizaram-se cinco lotes de sementes, e foram selecionadas 200 sementes aleatoriamente por lote, obtendo um total de 1000 sementes, as quais foram submetidas à análise do infravermelho próximo. As sementes foram analisadas individualmente colocadas em suporte circular do próprio equipamento e sobre este foi colocado um cilindro metálico para não ocorrer interferência dos feixes de luz e para que os resultados espectrais sejam mais precisos. Os dados espectrais de cada semente foram obtidos por meio do espectrômetro de transformada de Fourier – FT-NIR (Thermo Scientific Antaris II). Os espectros de refletância NIR foram gerados como $\log(1/R)$, onde o R é a refletância. Foram obtidos 3.112 pontos por espectro para cada semente, dentro de uma faixa de comprimento de onda que varia de 1000 a 2500 nm. Foram realizadas 200 varreduras contínuas dos espectros médios de cada lote, sendo feitas em 30 segundos para cada semente. Após a obtenção dos espectros, as sementes de cada lote foram submetidas ao teste de germinação com quatro repetições de 50 sementes, colocadas em papel toalha (Germitest®) umedecidos com água destilada, na quantidade de 2,5 vezes o peso do papel seco. Após a semeadura, foram confeccionados rolos e mantidos em germinador à temperatura de 25 °C. As avaliações foram realizadas no quinto e no décimo quarto dia após a semeadura e os resultados foram expressos em percentagem de plântulas normais, anormais e sementes mortas (BRASIL, 2009).

3.3.2 Delineamento experimental e análise estatística

O teste de germinação foi conduzido em delineamento inteiramente casualizado (DIC), com cinco lotes e quatro repetições. Os dados foram testados quanto à distribuição normal dos erros pelo teste de Shapiro-Wilk e à homogeneidade de variâncias pelo teste de Bartlett. Os dados foram submetidos à análise de variância e as médias obtidas para cada lote foram comparadas pelo teste de Tukey a 5% de probabilidade.

Os espectros NIR obtidos foram classificados conforme os resultados do teste de germinação, avaliando o desempenho fisiológico das sementes em alto, intermediário e baixo vigor. Para uma avaliação inicial exploratória, foram gerados gráficos dos espectros originais, acompanhado da análise de componentes principais (PCA) (JOLLIFFE; CADIMA, 2016).

Foram gerados gráficos da média dos espectros obtendo-se um único espectro para cada classe. Em seguida, os dados foram submetidos aos métodos de pré-processamentos e a construção de modelos de classificação. Foram realizados pré-processamentos como: centralização, correção do espalhamento multiplicativo (MSC), variação normal padrão (SNV), 1º e 2º derivada pelo método Savitzky-Golay, utilizando-se janela de 11 variáveis (SAVITZKY; GOLAY, 1964). O número de variáveis latentes (VL) foi determinado por validação cruzada aleatória com dez divisões.

Para cada combinação de pré-processamento, foi realizado um modelo de classificação através da análise de regressão por mínimos quadrados parciais (PLS-DA) (BARKER; RAYENS, 2003), onde 70% dos dados foram utilizados para treinamento e 30% utilizados para validação. O comportamento dos modelos foi avaliado pela acurácia e medida de confiabilidade kappa.

Foram identificadas as faixas de comprimento de onda que mais contribuíram para a geração do modelo das diferentes classes de vigor.

Para todas as análises dos dados, foi utilizado o software estatístico R 4.1.1. (R CORE TEAM, 2020), com auxílio dos pacotes Nira, Caret e Prospect.

4. RESULTADOS E DISCUSSÃO

4.1. Caracterização da qualidade inicial dos lotes de arroz

O grau de umidade das sementes dos diferentes lotes de arroz variou de 13,3 a 13,6% (Tabela 1). Segundo Marcos-Filho et al. (2015) e Worma et al. (2019) a uniformização da umidade, com variação menor que 0,5%, é importante para se obter resultados confiáveis e consistentes de qualidade fisiológica dos lotes.

Ao avaliar a germinação (Tabela 1), verifica-se que os lotes 1, 2, 3 e 4 não difeririam entre si e foram superiores ao lote 5. Entretanto é importante observar que a germinação de todos os lotes está dentro do padrão mínimo exigido para a comercialização de sementes de arroz no Brasil, que é de 80% (Brasil, 2012). Segundo Queiroz et al. (2019), este é um teste

onde as sementes podem expressar seu máximo potencial fisiológico já que são submetidas a condições ideais de umidade, luz e temperatura. Resultados semelhantes foram obtidos na avaliação da primeira contagem de germinação, onde se verifica que o lote 5 foi o de pior desempenho. Este teste complementa os resultados de germinação, e pode ser interessante para a classificação dos lotes quanto ao vigor, uma vez que fornece informações relacionadas à velocidade de germinação dos diferentes lotes (NAKAGAWA, 2020). Mas no presente estudo não se mostrou eficiente para esta finalidade.

Tabela 1 – Caracterização da qualidade fisiológica de cinco lotes de sementes de arroz cv. IRGA 424 avaliados pelo grau de umidade (GU), germinação (G), primeira contagem de germinação (PCG), envelhecimento acelerado (EA), índice de velocidade de emergência (IVE) e germinação a baixa temperatura (GBT).

Lotes	GU(%)	G (%)	PCG(%)	EA (%)	IVE (índice)	GBT(%)
1	13,4	96 a	86 a	91 a	91 a	84 a
2	13,6	92 a	85 a	81 b	91 a	83 a
3	13,3	96 a	83 a	77 c	92 a	79 ab
4	13,4	92 a	84 a	72 c	88 b	77 b
5	13,4	85 b	74 b	59 d	85 b	64 c
Fc	-	6,32*	6,51*	112,02 *	10,13 *	20,90 *
CV (%)	-	3,96	4,51	2,97	1,86	4,54

* = significativo pelo teste F a 5% de probabilidade; F = valor de F calculado; CV = coeficiente de variação. Médias seguidas pela mesma letra, na coluna, não diferem entre si pelo teste de Tukey, a 5% de probabilidade.

Pode-se observar que pelo teste de envelhecimento acelerado foi possível classificar os lotes em níveis de vigor, com maior valor para o lote 1, seguido pelos lotes 2, 3 e 4, com nível intermediário de vigor, sendo o lote 5 inferior aos demais (Tabela 1). Este é um teste que permite avaliar o potencial fisiológico das sementes quando a taxa de deterioração das sementes aumenta pela exposição a altas temperaturas e umidade relativa (BAALBAKI, 2009; MARCOS-FILHO, 2009); sementes de baixo vigor terão baixa germinação após esse estresse. É um teste que fornece uma estimativa do potencial de armazenamento das sementes e do potencial de emergência em campo. Observa-se, portanto, que sob condições ideais do teste de germinação, foi possível separar apenas o lote 5 como inferior aos demais. Já pelo teste de envelhecimento acelerado, que é um teste de estresse, houve uma maior estratificação dos lotes em níveis de qualidade.

Os resultados do índice de velocidade de emergência (IVE) apresentam os lotes 1, 2 e 3 com melhor desempenho, diferindo estatisticamente dos lotes 4 e 5 (Tabela 1). Este teste permite separar lotes que germinam rápido daqueles que tem emergência mais lenta, característica importante para o desempenho no campo, uma vez que lotes com germinação mais rápida têm maior possibilidade de escapar de condições adversas em campo.

No teste de germinação a baixa temperatura, os lotes 1 e 2 apresentaram melhor desempenho e diferiram dos lotes 4 e 5 estatisticamente, lotes estes que também estiveram entre os de pior desempenho nos demais testes de vigor utilizados. Destaca-se o lote 5, com desempenho inferior aos demais assim como observado no teste de germinação, primeira contagem e envelhecimento acelerado (Tabela 1). Segundo Grzybowski et al. (2015), as sementes de alta qualidade quando são submetidas a estes estresses abióticos apresentam melhor desempenho.

Em síntese, com base nos testes utilizados para a caracterização inicial dos diferentes lotes de arroz quanto a seu potencial fisiológico, observa-se que os lotes 1 e 2 estiveram entre os de melhor potencial fisiológico em todos os testes aplicados, exceto no de envelhecimento acelerado. Apenas no teste de envelhecimento acelerado (Tabela 1), o lote 5 apresentou pior desempenho. O uso de lotes com diferenças na qualidade fisiológica é importante para a identificação dos níveis de vigor de sementes, especialmente quando se avalia ou define com técnicas inovadoras que estão sendo testadas para tal fim de acordo com cada espécie.

4.1. Ensaio I – Análise computadorizada de plântulas pelos softwares SAPL® e ILASTIK para avaliação do potencial

4.2.1. Análise de plântulas pelo software SAPL®

Os resultados obtidos através da análise computadorizada de plântulas pelo processamento de imagens com o software SAPL® encontram-se na Tabela 2. Nas avaliações feitas aos 4 dias após o início do teste de germinação, observa-se que pelo CPA o lote 1 obteve o melhor desempenho, seguido pelo lote 2, sendo ambos superiores aos lotes 3 e 4 que não diferiram entre si, enquanto o lote 5 foi inferior aos demais, com o menor crescimento da parte aérea. Esta classificação dos lotes quanto ao potencial fisiológico foi semelhante à obtida no teste de envelhecimento acelerado (Tabela 1) que é um teste de vigor confiável e bastante utilizado para sementes de diferentes culturas, inclusive arroz (Marcos-Filho et al., 2020). Desta

forma, pode-se afirmar que a avaliação do comprimento da parte aérea das plântulas de arroz aos quatro dias após a semeadura pode se constituir em alternativa interessante para a avaliação do vigor das sementes, já que fornece resultados relativamente rápidos, principalmente em relação aos de envelhecimento acelerado. Este teste envolve a aplicação de estresse por calor e umidade nas sementes por um período de 72 h e mais cinco dias de germinação para que se obtenha os resultados, ou seja, um total de oito dias. Já para as avaliações de CPA feitas aos cinco e sete dias as diferenças entre os lotes não foi tão evidente, especialmente os de vigor intermediário, sendo possível apenas diferenciar os lotes de maior e menor desempenho. Lorentz e Nunes (2013) identificaram a qualidade fisiológica das sementes de arroz através de medições realizadas com régua do CPA aos quatorze dias após a semeadura. Porém, é possível otimizar o processo e reduzir o período de avaliação do comprimento da parte aérea de plântulas. O grau de umidade das sementes dos diferentes lotes de arroz variou de 13,3 a 13,6% (Tabela 1). Segundo Marcos-Filho et al. (2015) e Worma et al. (2019) a uniformização da de arroz para cinco dias através do processamento digital de imagens (CUADRADO, 2020; BRUNES et al., 2019).

Para o comprimento radicular (CR), no quarto dia de avaliação tem-se os lotes 1 e 2 como superiores aos demais, seguido pelo lote 3 que não diferiu do lote 4, mas foi superior ao lote 5, este lote teve novamente o pior desempenho, mas não diferiu do lote 4 (Tabela 2). Assim, pode-se afirmar que os lotes 1, 2 e 3 foram superiores ao lote 5. Resultados semelhantes foram obtidos aos 5 dias, apenas destacando que o comprimento radicular do lote 1 foi superior ao do lote 2. Aos sete dias, os resultados foram semelhantes aos obtidos aos 4 e 5 dias, mas nota-se que o lote 4 que vinha se igualando ao lote 5 de menor desempenho, não diferiu do lote 1, sempre indicado entre os de melhor desempenho. Assim, o CR avaliado principalmente no quinto dia após a semeadura apresentou resultados coerentes com os obtidos no teste de envelhecimento acelerado (Tabela 1). Em sementes de milho, Dias et al. (2015) ao avaliarem o CR, encontraram resultados similares entre os testes tradicionais de vigor e os obtidos por meio da análises de imagem.

Os resultados observados na Tabela 2, para o comprimento total de plântulas (CTP) indicam uma classificação dos lotes em níveis de vigor, aos 4, 5 e 7 dias. No entanto, aos quatro e 5 dias, obteve-se uma separação mais eficiente dos lotes quanto ao vigor, destacando o lote 1 como superior aos demais, seguido pelo lote 2, que foi superior ao lote 3, sendo estes lotes superiores ao lote 5. Pasqualli (2005) avaliando pelo método manual através do comprimento de plântulas aos sete dias não mostrou diferenças entre lotes de arroz em relação ao vigor, isto devido ao alto potencial fisiológico que podem apresentar os lotes. A relação entre comprimento

de plântulas e o vigor das sementes possibilita uma estimativa mais eficaz para o controle de qualidade fisiológica dos lotes (NAKAGAWA et al., 1999). Segundo Krzyzanowsky et al. (2020), sementes vigorosas desenvolvem plântulas mais uniformes e fortes, capazes de produzir de forma mais rápida mecanismos de reparo, sínteses de novos tecidos e mobilização de reservas durante o processo de germinação.

Em geral, verifica-se que nas avaliações feitas no quarto dia, o CPA e CTP (Tabela 2) foram eficientes para classificar os lotes em níveis de vigor de modo similar ao obtido no teste de envelhecimento acelerado (Tabela 1). Já para o CR, informações mais relacionadas aos testes de vigor relacionados na Tabela 1, principalmente envelhecimento acelerado, foram obtidas no quinto dia. Cuadrado (2020) observou que pelo SAPL, aos quatro e cinco dias, foi possível classificar lotes de sementes de arroz cv-Relâmpago utilizando o CR. Em sementes de milho, a diferenciação dos lotes quanto ao vigor foi obtida pelo software SAPL quanto se utilizou o CTP aos três dias (SILVA, 2020). Já Senna et al. (2015) avaliando o CPA por meio de medições com régua obteve resultados consistentes relacionados ao vigor aos sete dias após a semeadura.

Com relação aos índices gerados pelo SAPL, apenas o IU aos cinco dias e IV aos sete dias permitiram uma maior estratificação dos lotes quanto ao vigor. O IU forneceu informações similares às obtidas nas avaliações do CPA e CPT aos 4 dias e CR aos 5 dias (Tabela 2). Verifica-se ainda que o IU determinado aos 4 dias não permitiu a separação dos lotes quanto à qualidade fisiológica, não havendo diferença estatística entre eles. Aos sete dias, os resultados obtidos foram compatíveis com os do teste de germinação e primeira contagem (Tabela 1) destacando apenas o lote 5 como inferior aos demais que não diferiram entre si. Resultados semelhantes aos de germinação também foram observados para o IC aos quatro e setedias e IV aos quatro e cinco dias, onde observa-se diferença significativa para o lote 5, que apresentou menor qualidade. Os índices de IC e IV fornecidos pelo SAPL® mostraram sensibilidade para diferenciar o aspecto fisiológico de lotes de sementes de arroz, lentilha e moringa (CUADRADO, 2020; PEREIRA et al., 2020; LIMA O, 2022). Segundo Pereira et al. (2020) e Bennet (2002), estes índices podem ser importantes para a predição do vigor dos lotes de sementes.

O índice de uniformidade é um parâmetro importante relacionado ao potencial fisiológico das sementes, uma vez que a emergência das plântulas não homogênea pode retardar o desenvolvimento e o crescimento do estande nos estádios fenológicos iniciais pós-emergência, podendo resultar em desuniformidade nas fases de florescimento e maturação prejudicando a qualidade da semente na colheita (MARCOS-FILHO, 2015a).

1	0,66 a	3,11 a	3,77 a	882,94 a	240,27 a	366,75 a
2	0,53 b	2,60 a	3,14 b	876,18 a	216,76 a	353,52 ab
3	0,44 c	1,83 b	2,27 c	766,86 a	201,54 a	326,59 ab
4	0,43 c	1,47 bc	1,90 cd	682,28 a	212,16 a	328,15 ab
5	0,38 d	1,18 c	1,56 d	664,97 a	165,75 b	289,6 b
Fc	51,57*	36,00*	46,18*	3,54 ^{ns}	1,21 *	3,35*
CV (%)	6,35	7,1	9,6	7,16	9,68	8,72
5 dias						
1	1,71 a	5,96 a	7,67 a	757,04 a	520,56 a	579,51 a
2	1,70 a	5,87 b	7,01 b	734,07 ab	506,35 a	572,86 a
3	1,54 ab	5,32 bc	6,85 bc	727,25 b	507,72 a	517,53 a
4	1,51 ab	5,07 bc	6,57 bc	726,91 b	504,45 a	418,69 a
5	1,31 b	4,04 c	5,35 c	680,55 c	500,71 a	475,16 b
Fc	5,10*	3,47*	4,84 *	19,37*	2,71 ^{ns}	3,25*
CV (%)	8,36	9,6	5,48	1,74	7,12	9,03
7 dias						
1	3,38 a	7,24 a	10,63 a	857,89 a	672,62 a	739,10 a
2	3,07 ab	6,80 ab	9,87 ab	846,1 a	659,62 a	689,66 ab
3	3,08 ab	5,98 bc	9,06 bc	832,82 a	609,52 ab	666,01 bc
4	3,00 ab	6,30 ab	9,30 b	838,04 a	609,03 ab	640,98 bc
5	2,27 b	5,61 c	7,89 c	722,39 b	530,28 b	612,91 c
Fc	7,37 *	9,37 *	11,70*	5,35*	4,08*	11,29*
CV (%)	4,23	6,6	6,35	5,81	8,99	4,28

* = significativo pelo teste F a 5% de probabilidade; F = corresponde ao valor de F calculado; CV = coeficiente de variação. Médias seguidas pela mesma letra, na coluna, não diferem entre si pelo teste de Tukey, a 5% de probabilidade.

4.2.2. Análise de plântulas pelo software ILASTIK

Na Figura 3 podemos observar os resultados fornecidos pelo software ILASTIK nas avaliações feitas aos 4, 5 e 7 dias após da semeadura referentes à probabilidade de os lotes gerarem plântulas normais fortes, plântulas normais fracas, plântulas anormais e sementes mortas. Aos quatro dias, os lotes 1, 2 e 3 apresentaram alta probabilidade de gerarem plântulas normais fortes, não diferindo entre si, sendo superiores ao lote 5. Assim, foi possível separar os lotes em duas classes de vigor. Nota-se que os lotes 4 e 5 foram os que apresentaram maiores valores para sementes não germinadas em relação aos demais.

Já aos cinco dias, foi possível obter maior sensibilidade na distinção entre os lotes, observando-se maior vigor para as sementes dos lotes 1 e 2 em relação aos demais lotes. O lote 2 foi superior ao lote 3 que por sua vez foi superior ao lote 5, classificado como o de menor potencial fisiológico (Figura 3). Assim, verifica-se que houve uma separação bem definida entre os lotes quanto ao vigor, baseando-se na maior probabilidade de gerar plântulas vigorosas e menor probabilidade de gerar plântulas não vigorosas e sementes mortas. Medeiros e Pereira (2018) relatam que plântulas normais fortes as quais apresentam todas estruturas desenvolvidas, tem uma diferença quanto ao tamanho, vigor e a taxa de crescimento. Este é um aspecto importante porque plântulas fortes podem apresentar em campo menos sensibilidade a estresses e maior probabilidade de se obter estande adequado e uniforme.

Aos sete dias, também foi possível observar diferenças significativas entre os lotes quanto ao vigor, embora não tão detalhadas como as observadas aos cinco dias, uma vez que o lote 3 se igualou aos lotes 1 e 2, sendo classificados como os de maior vigor em comparação aos demais. O lote 5 foi o de pior desempenho assim como constatado aos cinco dias. Assim, aos sete dias obteve-se uma classificação dos lotes em três níveis de vigor pelos dados de PNF, enquanto com cinco dias, ou seja, com menor tempo, foram obtidos quatro níveis de vigor. Os dados referentes à PNFra, PA e SNG confirmam estas informações, uma vez que quanto menor a proporção destes três parâmetros maior o vigor do lote.

A avaliação do potencial fisiológico das sementes por meio de softwares de código aberto como o ILASTIK é de grande relevância para o setor de produção de sementes pois permite uma avaliação criteriosa e sem morosidade. Para Medeiros et al. (2020a), a aplicação de aprendizagem de máquina com dados gerados pelo software ILASTIK para a classificação da qualidade fisiológica de sementes e plântulas de soja é uma ferramenta precisa, pois permitiu identificar as sementes quanto aos danos e classificar as plântulas quanto ao vigor. Cuadrado em (2020) em sementes de arroz e Limão (2022) em sementes de lentilha, relatam a eficiência do software ILASTIK, para a classificação do potencial fisiológico dos lotes correlacionados com resultados similares aos testes de vigor utilizados. Os resultados obtidos com o software ILASTIK em plântulas de arroz da cultivar Relâmpago aos quatro e cinco dias após o início do teste de germinação permitiram a estratificação dos lotes em níveis de vigor fornecendo informações coerentes a classificação obtida nos testes de vigor tradicionais utilizados (CUADRADO, 2020).

O vigor compreende o conjunto de características que incluem uniformidade de germinação, desenvolvimento inicial das plântulas, rapidez na emergência sob diferentes condições de campo, dentre outras (ISTA, 2019). Assim, a classificação do vigor das plântulas

é uma característica importante, que pode ser facilmente obtida durante a avaliação do teste de germinação, sem a necessidade de condução de testes adicionais com metodologias mais complexas ou que demandam maior tempo para a obtenção dos resultados. Neste contexto, tecnologias que permitam identificar sementes capazes de gerar plântulas vigorosas ou não, são de grande relevância para a avaliação rápida e segura do potencial fisiológico dos lotes, agilizando as tomadas de decisões sobre o destino dos lotes a serem comercializados.

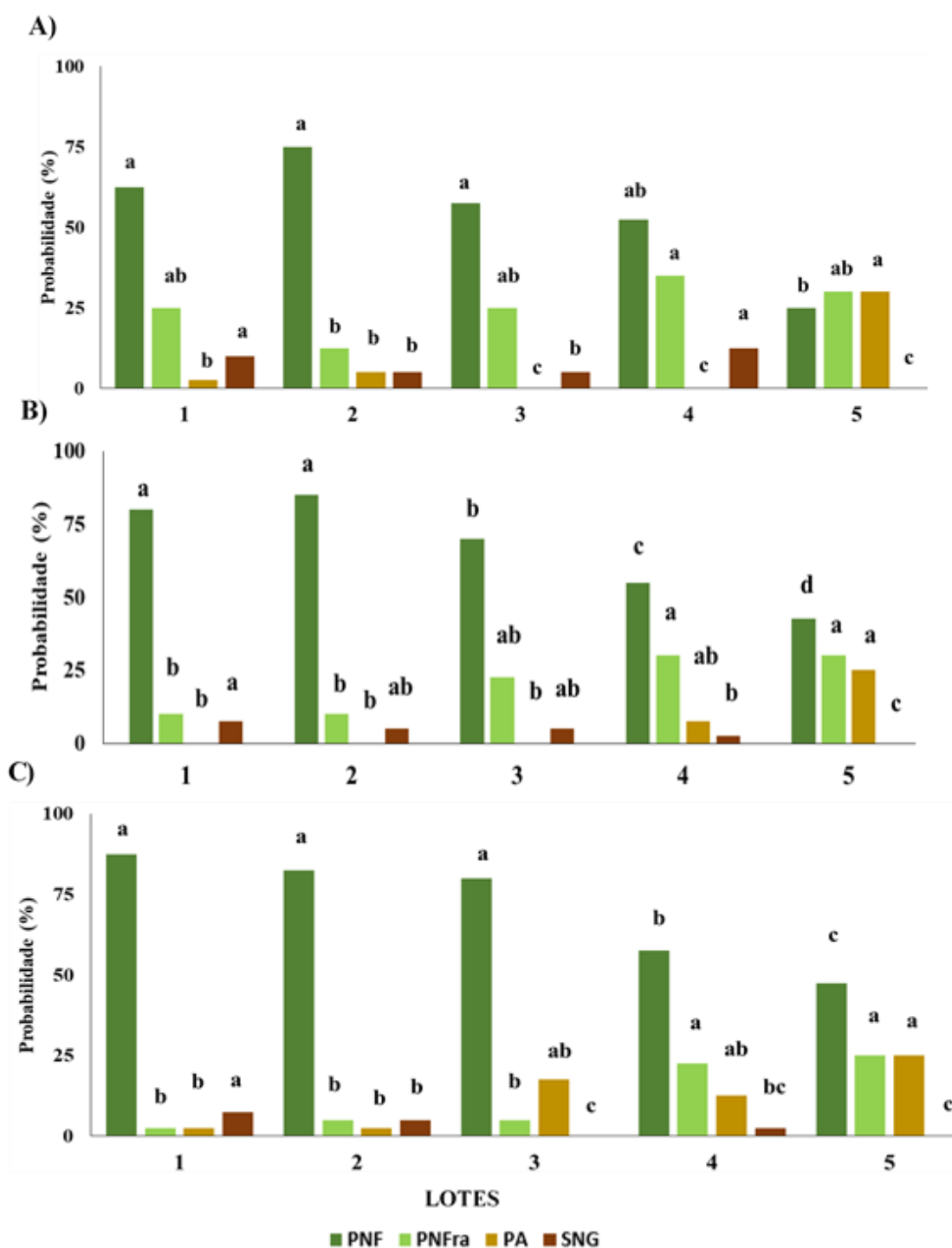


Figura 3. Probabilidade de os lotes de sementes gerarem plântulas normais fortes (PNF), plântulas normais fracas (PNFra), plântulas anormais (PA) e sementes não germinadas (SNG) no 4º (A), 5º (B) e 7º (C) dia após a semeadura, de acordo com dados obtidos pelo software ILASTIK.

Estudos com diferentes espécies como milho e soja (MEDEIROS et al., 2019), feijão (GOMES et al., 2014), trigo (BRUNES et al., 2016), entre outros, têm indicado que a utilização

de sistemas computacionais é adequada para a classificação de lotes em menor tempo em relação aos testes tradicionais de vigor.

Segundo Limão (2022) os modelos desenvolvidos no software ILASTIK podem ser testados para melhorar e permitir obter resultados mais precisos, através do acréscimo da quantidade de imagens, lotes e tratamentos. Neste contexto, o software ILASTIK foi adequado na avaliação da qualidade fisiológica dos lotes de arroz, sendo uma alternativa rápida, de fácil acesso e manipulação.

4.2. Ensaio II – Espectroscopia no infravermelho próximo (NIR) e sua relação com o potencial fisiológico das sementes

Inicialmente, foi determinado o grau de umidade das sementes de todos lotes antes da condução deste estudo, e os valores obtidos foram uniformes, ou seja, em torno de 12%. Em seguida, foi realizado o teste de germinação para a classificação dos lotes quanto ao potencial fisiológico (Figura 4). Considerando o padrão mínimo de germinação para comercialização de sementes de arroz, o lote que apresentou germinação abaixo de 80% foi classificado como de baixa qualidade. Lotes que apresentaram porcentagem igual ou superior a 90% foram classificados como de alta qualidade (lotes, 1, 2 e 3). O lote 4, com germinação muito próxima ao limite mínimo para comercialização, foi considerado como de qualidade intermediária.

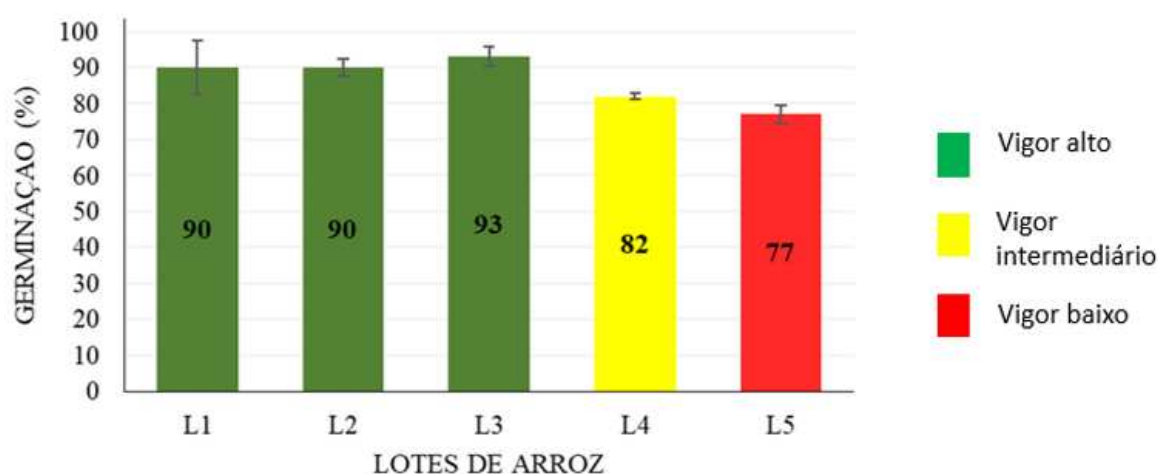


Figura 4. Germinação e classificação da qualidade fisiológica de cinco lotes de arroz cv. IRGA 424.

O teste de germinação é o método mais frequentemente utilizado para a avaliação da qualidade fisiológica das sementes, sendo referência para se aceitar ou recusar lotes para a comercialização. Cabe ressaltar que o teste de germinação para sementes de arroz compreende um período de 14 dias para que os resultados sejam obtidos. Por isso, a implementação de novas metodologias que consigam prever o potencial fisiológico das sementes e fazer uma pré-seleção dos lotes de baixa qualidade pode otimizar a agilidade dos processos internos de tomadas de decisão nas empresas quanto ao destino dos lotes.

Por meio da Figura 5A, é possível observar os 1000 espectros originais, seguido pela média dos espectros originais (Figura 5B) de acordo com a classificação da qualidade fisiológica obtida pelo teste de germinação, onde já é possível observar uma tendência de separação das classes quanto ao potencial fisiológico.

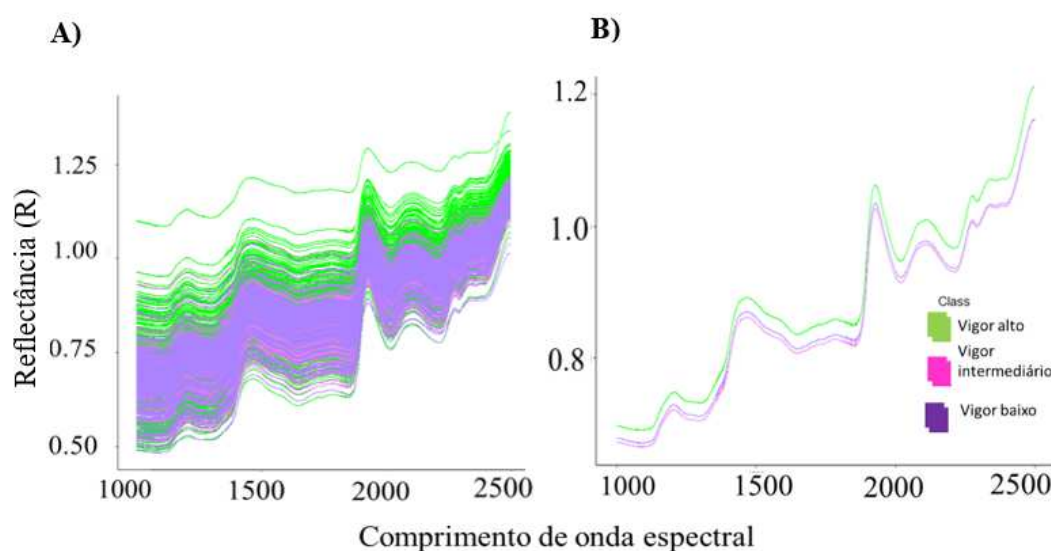


Figura 5. Espectros FT-NIR de diferentes lotes de sementes de arroz cv. IRGA 424: Espectros originais (A) espectros obtidos para cada classes de qualidade de semente (B) espectros médios originais.

É possível observar na Figura 5A um agrupamento dos espectros totais das sementes de alto vigor diferenciando-se dos espectros das sementes de baixo vigor. Já para as médias dos espectros esta tendência foi mantida em que o espectro das sementes de alto vigor continua diferenciando-se dos espectros das de médio e baixo vigor. A representação dos espectros pode ser usada para calcular compostos específicos, como açúcares, proteínas e teor de água das sementes (BADARÓ et al., 2019), como também para a classificação de sementes de diferentes

variedades ou genótipos, sempre que estes sejam influenciados pela composição química (COZZOLINO, 2016).

Análise de componentes principais (PCA) dos espectros originais, explicaram mais de 99% da variação destes, sendo possível identificar previamente uma tendência de separação das classes de potencial fisiológico, principalmente para os lotes de alto desempenho (Figura 6).

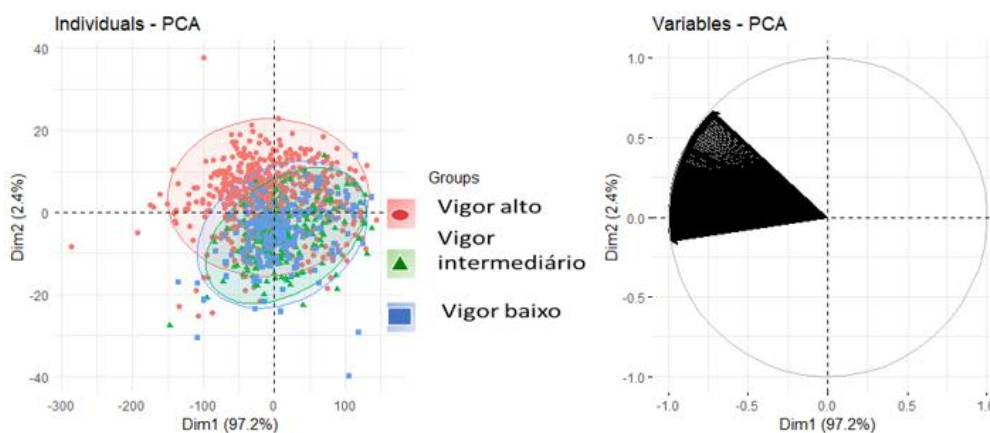


Figura 6. Análise de componentes principais usando os dados originais FT-NIR obtidos de sementes de diferentes lotes de arroz cv. IRGA 424. Diagrama de ordenação central (A) e círculo de correlação e distribuição (B).

A representação bidimensional mostra um agrupamento das pontuações vermelhas correspondentes aos lotes de alto vigor. No entanto, para os lotes de vigor intermediário e baixo vigor, houve uma sobreposição das pontuações verdes e azuis, de modo que não foi possível obter uma separação efetiva de cada classe. Assim, verifica-se que foi possível destacar apenas os lotes de alto potencial fisiológico dos demais (Figura 6A). Limão (2020) observou que a análise de componentes principais dos espectros NIR não foram eficientes ao ponto de rastrear diferentes classes de sementes de lentilha quanto ao vigor. Resultados semelhantes foram obtidos por Ribeiro et al. (2020) em sementes de grão-de-bico. Porém, as análises espectrais necessitam de processamento dos dados com a finalidade de aprimorar o modelo de classificação e remover fenômenos de ruídos físicos nos espectros, como espalhamentos e desvios da linha de base por exemplo (RINNAN et al., 2009).

Uma vez obtidos os espectros originais foram utilizados então diferentes tipos de pré-processamentos: centralização, variação normal padrão (SNV), correção do espalhamento multiplicativo (MSC) e a derivada de Savitzky-Golay (SG). Através do método de classificação PLS-DA foram gerados modelos de separação dos três níveis de qualidade fisiológica dos lotes de arroz (Tabela 3). Os diferentes tipos de pré-processamentos dos espectros podem aprimorar

a capacidade preditiva ou discriminatória do modelo ao remover variáveis espectrais que podem alterar informações importantes para a análise dos dados (PASQUINI, 2018). No entanto, é importante escolher o método de pré-processamento mais correto para não remover informações importantes. Uma vez analisados os espectros originais com os diferentes tipos de pré-processamento, os resultados do modelo devem manter ou diminuir a complexidade do modelo (RINNAN et al., 2009). Segundo Lee (2011), o conjunto de treinamento é utilizado para desenvolver os modelos de classificação e encontrar parâmetros ideais para a validação. O conjunto de validação refere-se à utilização de amostras que não foram utilizadas no treinamento, para assim fornecer uma estimativa do comportamento da equação de classificação analisada em futuras amostras.

Tabela 3. Resultados de acurácia e confiabilidade (kappa) para o conjunto de dados de treinamento e validação (teste) dos modelos de classificação gerados por meio da análise PLS-DA com diferentes tipos de pré-processamentos em espectros FT-NIR de sementes de cinco lotes de arroz cv. IRGA 424.

Pré-processamento	Treinamento (n=700)		Validação (n=300)	
	Acurácia	Kappa	Acurácia	Kappa
Espectros originais	0.91	0.84	0.91	0.85
Centralização	0.91	0.84	0.91	0.85
SNV	0.85	0.72	0.86	0.73
MSC	0.85	0.72	0.86	0.78
1º Derivada SG	0.98	0.97	0.94	0.90
2º Derivada SG	0.99	0.99	0.96	0.93

n = número de espectros usados para treinamento e/ou teste do modelo de classificação. SG = Savitzky-Golay; MSC = *Multiplicative Scatter Correction* (correção do espalhamento multiplicativo); SNV = *Standard Normal Variate* (variação normal padrão).

A combinação entre análises multivariadas e diferentes tratamentos de pré-processamentos são essenciais para aperfeiçoar a acurácia e a confiabilidade (coeficiente Kappa) do modelo com base nos espectros FT-NIR, pois os espectros resultam em várias combinações de tons e sobretons gerando ruídos e alterações nos resultados (WU et al., 2009). Segundo Ribeiro et al. (2020) e Baum et al. (2012), a utilização de diferentes técnicas de pré-processamentos são necessárias para minimizar os ruídos e otimizar a escolha do modelo, de modo a obter informações relativas à classificação eficientes e precisas.

Na tabela 3, observa-se que todos os modelos obtidos para os diferentes tipos de pré-processamento testados apresentaram alta acurácia e alta confiabilidade (kappa). O primeiro modelo classificatório baseado nos espectros originais, sem qualquer tipo de pré-processamento de dados junto com o método de pré-processamento de centralização demonstraram valores iguais onde a acurácia para o conjunto de treinamento e validação (0,91) do modelo e a confiabilidade foi de (0,84) para o conjunto de treinamento e 0,85 para validação. Para SNV e MSC valores iguais foram observados tanto para a acurácia (0,85) e kappa (0,72) para o conjunto de treinamento. Quando foi feita a validação os valores obtidos para SNV e MSC para a acurácia também foram iguais, ou seja, 0,86. Enquanto o kappa foi menor para o SNV (0,73) em comparação ao obtido com o MSC que foi de 0,78. Na determinação do modelo de classificação, o coeficiente kappa indica a aceitação e a credibilidade do modelo (COHEN, 1968).

Com a aplicação dos pré-processamentos com a 1ª e 2ª derivada de Savitzky-Golay, obteve-se um incremento nos valores de acurácia e kappa (Tabela 3). Por meio da 1ª derivada de SG, os valores no conjunto de treinamento para a acurácia foram 0,98 e 0,97 para kappa; já para a validação do modelo a acurácia foi de 0,94 e kappa de 0,90. Assim, foi possível observar que a utilização do modelo de classificação de PLS-DA com a 2ª derivada pelo filtro de Savitzky-Golay foi a combinação mais eficiente para a classificação dos lotes de sementes de arroz quanto ao potencial fisiológico, onde para o treinamento os valores da acurácia e coeficiente kappa obtidos foram mais próximos de 1,0, uma vez que para o conjunto de validação a acurácia foi de 0,96 e o kappa 0,93 (Tabela 3). Em estudos recentes com sementes de soja (AMANAH et al., 2020), feijão preto (YANG et al., 2015), trigo (FAN et al., 2020), uva (ZHAO et al., 2018) e milho (AMBROSE et al., 2016) foram obtidas acurácias superiores a 90%. Diversos autores utilizaram a derivada de Savitzky-Golay para a suavização dos ruídos nos espectros originais em estudos com sementes de lentilha (LIMAO, 2022), soja (SILVA, 2020; LEE e CHOUNG, 2011) e grão-de-bico (RIBEIRO et al., 2020). Limão (2022) constatou que foi possível confirmar o modelo de classificação de PLS-DA a partir da primeira derivada de Savitzky-Golay como o mais eficiente para a classificação do vigor de lotes de sementes de lentilha.

No entanto, evidenciamos que os valores da acurácia e kappa tanto para treinamento como para validação dos pré-processamentos SNV e MSC foram inferiores aos obtidos com os espectros originais (Tabela 3). Podemos associar esses resultados com a aparição mínima desses efeitos nos espectros originais. Outro fator que pode ter contribuído para a redução na eficiência de classificação é que os métodos de correção e dispersão (MSC e SNV), removeram

informações importantes para a separação das classes (RINNAN et al. 2009). Segundo Folha (2014), o método MSC é com frequência utilizado para a correção do efeito da dispersão da luz presentes nos espectros de infravermelho próximo que utilizam técnicas de refletância como é o caso do FT-NIR utilizado na presente pesquisa. O método SNV apresenta um efeito similar ao MSC, sendo utilizado para a correção das interferências de espalhamento e os efeitos do espelhamento entre diferentes classes de espectros (SANTOS, 2007).

No presente estudo, com base no modelo em que foram utilizados os pré-tratamentos da 2ª derivada de Savitzky-Golay e através do gráfico de pontuação da PLS-DA, foi possível observar um agrupamento diferenciado dos lotes de alto potencial fisiológico em comparação aos lotes das demais classes (qualidade intermediária e baixa). (Figura 7). Para a construção dos modelos classificatórios, a técnica PLS-DA é mais comumente utilizada, possibilitando a separação para o conjunto de amostras em estudo (BRERETON; LLOYD, 2014). Segundo Venkatesan et al. (2020), na área de sementes a PLS-DA é o método mais eficiente para a classificação das sementes quanto à sua qualidade.

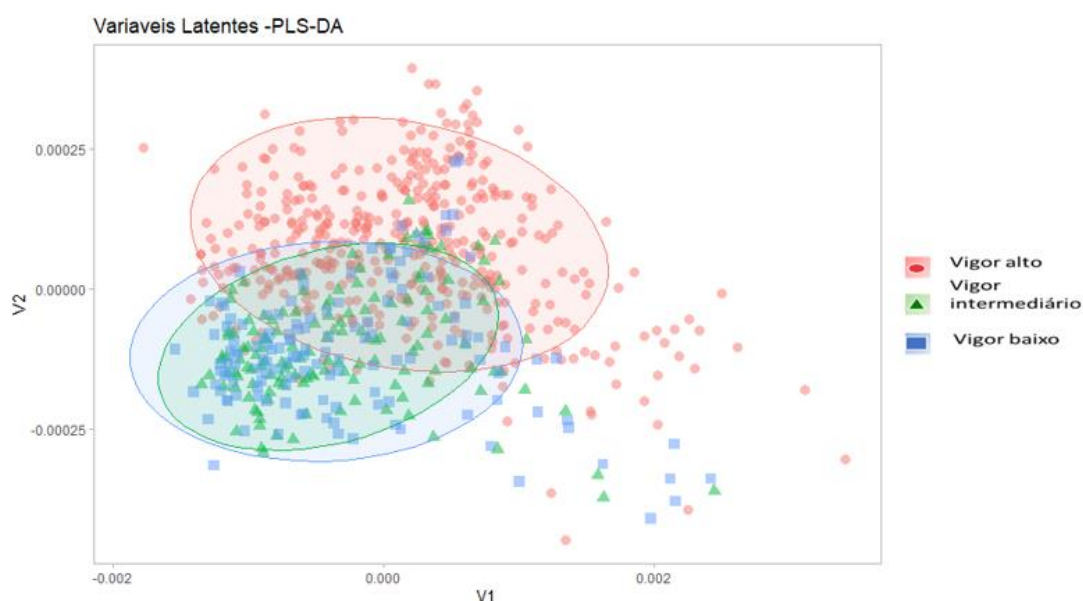


Figura 7. Gráfico de pontuação da análise discriminante por mínimos quadrados parciais (PLS-DA) com os dois primeiros componentes discriminantes de sementes de arroz com base nos espectros FT-NIR transformados pela segunda derivada Savitzky-Golay. As elipses mostram intervalos de confiança de 95% em cada classe de sementes.

Soares et al. (2016) utilizando a técnica NIR para avaliação de qualidade de sementes de algodão de quatro cultivares e comparando a eficiência de dois métodos de classificação, o SPA-LDA (Algoritmos das Projeções Sucessivas associado com o modelo de Análise Discriminante Linear) e o PLS-DA, obtiveram 97% de sementes classificadas adequadamente na

primeira validação. Sementes de diferentes espécies de Brassicas foram separadas com base no teor de óleo com acerto de 100% utilizando o modelo PLS-DA (MEDEIROS, 2021). Outros estudos para avaliação do vigor em sementes de milho doce (WANG et al., 2020) e para a detecção da viabilidade em sementes de soja (KUSUMANINGRUM et al., 2018), registraram precisão da classificação superior a 98% com o método PLS-DA.

A Figura 8 representa a matriz de confusão após o pré-processamento com a segunda derivada de Savitzky-Golay com base na classificação dos lotes de sementes de arroz. É possível observar para o conjunto de treinamento que as classificações de alto e médio vigor obtiveram um acerto de 100%, no entanto só três sementes não foram classificadas na categoria de baixo potencial fisiológico.

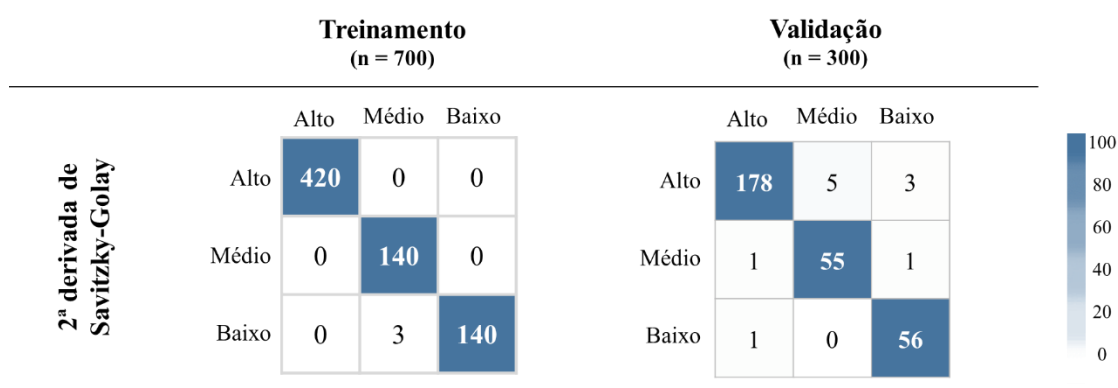


Figura 8. Matriz de confusão da classificação dos lotes de sementes de arroz cv. IRGA 424 quanto ao potencial fisiológico após o pré-processamento com a segunda derivada de Savitzky-Golay.

Para o conjunto de validação, que corresponde ao 30% das amostras que não foram processadas no conjunto de treinamento, observa-se, para a classe de alto potencial fisiológico, que de um total de 186 sementes somente oito não foram classificadas nesta classe, ou seja, 5 classificadas na categoria de médio vigor e 3 na de baixo vigor. Já para a classe de médio e baixo potencial fisiológico, o número de sementes que não foram incluídas em suas respectivas classes foi de duas e uma. Em geral, pode-se observar uma ocorrência mínima de erros para as três classes nos dois conjuntos, treinamento e validação. Estas informações permitem identificar a porcentagem de acerto por meio da suavização da segunda derivada de Savitzky-Golay, permitindo obter resultados precisos e confiáveis quanto à classificação do potencial fisiológico dos diferentes lotes de arroz (Figura 8).

Com base no modelo de classificação PSL-DA foram construídos os gráficos da importância das variáveis para cada classe de potencial fisiológico (Figura 9). Pode-se observar que

a identificação de faixas foi similar para os três níveis de qualidade. Os comprimentos de onda nas faixas entre 1000 – 1200 nm, 1350 -1470 nm, 1800 – 1900 nm, 2300 e 2500 nm, tiveram maior importância para a distinção das classes de alto potencial fisiológico. Para as sementes com vigor intermediário, as regiões mais importantes compreenderam os comprimentos de onda entre 1000- 1200, 1700-1800 e 1700-1900. Nota-se que os comprimentos de onda entre 1000 e 1200, 1800 e 1900 e o comprimento de onda entre 2300- 2500 foram importantes para a classificação das sementes de baixo potencial fisiológico.

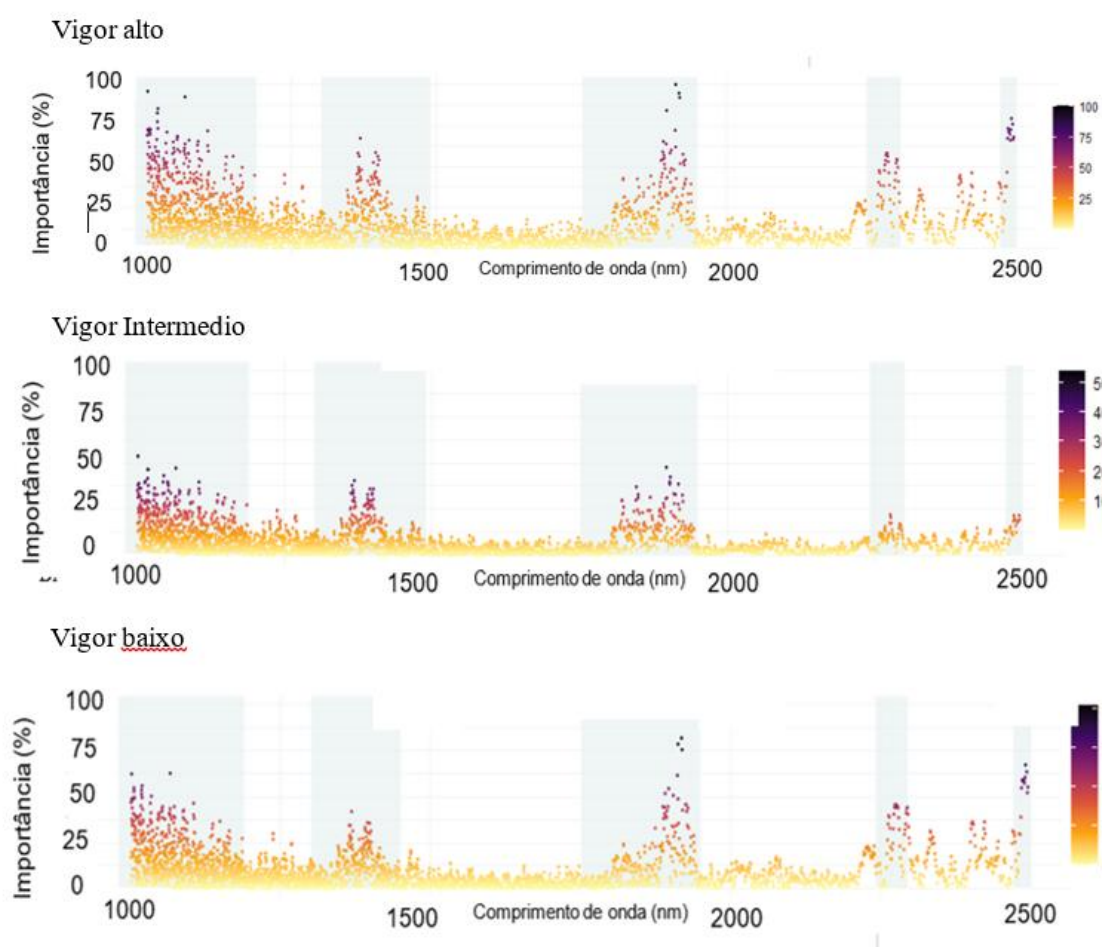


Figura 9. Importância das variáveis de comprimento de ondas para a classificação via PLS- DA dos níveis de vigor dos lotes de arroz cv. IRGA 424.

Para cada faixa da espectroscopia no infravermelho os comprimentos de absorção estão associados a diversos grupos funcionais relacionados com água, proteínas, lipídios, ácidos graxos e carboidratos, os quais se sobrepõem em várias zonas da faixa espectral (SHENK et al., 2007). A maioria dos componentes das sementes de arroz, como carboidratos, lipídios, proteína e umidade, contém uma grande quantidade de grupos de hidrogênio que são absorvidos pela espectroscopia no infravermelho próximo (CHEN et al., 2019).

Os comprimentos de onda na faixa 1000-1200 apresentaram maior agrupamento de pontuações, contribuindo junto com as outras faixas já mencionadas para a diferenciação da qualidade de sementes (Figura 9). Comprimentos de onda entre 1000-1200, representam a absorção de moléculas de sacarídeos (WORKMAN; WEYER, 2007); proteínas (XU et al., 2020); ácidos graxos (HOURANT et al., 2000); teor de óleo (MEDEIROS et al., 2022); carboidrato (SHENK et al., 2007).

Outra faixa importante para a contribuição da separação do potencial fisiológico das sementes foi na entre os comprimentos de onda 1800-1900 nm (Figura 9). Estas faixas representam a absorção de celulose (WORKMAN; WEYER, 2007); carboidratos associados com proteínas (KUSUMANINGRUM et al., 2018). Para os comprimentos de onda de 2300 e 2500 nota-se que para cada classe de vigor, tiveram uma contribuição importante para a estratificação quanto a qualidade fisiológica. Estes comprimentos estão relacionados com a absorção de carboidratos, proteínas, açúcares e celulose (WORKMAN; WEYER, 2007)

Em geral, pela Figura 9, observam-se diferentes faixas, com diferentes porcentagens de importância, as quais associadas à absorção de diferentes compostos químicos presentes nas sementes de arroz contribuíram para a classificação do potencial fisiológico. Cabe ressaltar que, compostos químicos como as proteínas, e os carboidratos estão presentes em porcentagens maiores nas sementes de arroz (MATSUO et al., 2018). E estes compostos podem estar relacionados diretamente com a qualidade fisiológica (DOMINGUES et al., 2012).

Para Delarmelino-Ferraresi et al. (2014) a maior quantidade de proteínas pode ser um indicativo de maior vigor nas sementes. Segundo Marcos-Filho (2020), durante o processo de deterioração das sementes ocorre uma série de manifestações bioquímicas como a diminuição da síntese de proteínas e a desnaturação das mesmas, resultando em declínio na qualidade fisiológica dos lotes. Para Peske et al. (2003), os carboidratos representam em média 85% da composição química média da semente de arroz. Segundo Erbas et al (2016), à medida que a semente se desenvolve, ocorre acúmulo de reservas como carboidratos, açúcares, amido e celulose, as quais são mobilizadas durante o processo de germinação, sendo vital para a retomada do crescimento do embrião, originando plântulas vigorosas. Portanto, espera-se que sementes com maior acúmulo de reservas como proteínas e carboidratos tenham maior potencial fisiológico e um melhor desempenho nas diferentes condições de ambiente (GU et al., 2016).

Em sementes de arroz, o conteúdo de proteínas permitiu diferenciar lotes em níveis de vigor e potencial de emergência em campo (BORTOLOTTTO et al., 2008). Resultados semelhantes foram obtidos com sementes de feijão (EHRHARDT-BROCARD; COELHO, 2022) e soja (HENNING et al., 2010). Para a cultura do arroz, a técnica NIR foi estudada e com base

em um modelo de calibração desenvolvido, permitiu a determinação do teor de amilose no elevador de grãos. (MATSUO et al., 2018). Em sementes de milho super- doce, as determinações de carboidratos e proteínas pelo FT NIR permitiu detectar diferenças entre sementes viáveis e não viáveis (QIU et al., 2018b).

De modo geral, os espectros NIR podem apresentar informações qualitativas e quantitativas de suma importância, com base nos compostos orgânicos das sementes, permitindo a identificação de diferenças entre amostras de maneira mais rápida em comparação a outras metodologias tradicionais (VENKATESAN et al., 2020).

Segundo Limão (2022), para o setor sementeiro a utilização da espectroscopia no infravermelho próximo para a classificação de sementes quanto à qualidade fisiológica, é importante por permitir rapidez na obtenção de resultados confiáveis e precisos, além de ser uma técnica não destrutiva. A metodologia utilizada neste estudo para a classificação dos lotes de sementes de arroz quanto ao potencial fisiológico através do FT NIR foi eficiente. No entanto, os estudos deverão ter continuidade para que técnica da espectroscopia no infravermelho próxima para a predição da qualidade fisiológica de sementes seja consolidada e possa ser utilizada nos programas de controle de qualidade das empresas produtoras de sementes.

5. CONCLUSÕES

A aplicação de tecnologias baseadas em análises de imagens de plântulas pelos softwares SAPL®, aos quatro dias, e ILASTIK, aos cinco dias, foi eficiente para a classificação do vigor de sementes de arroz, obtendo resultados rápidos e confiáveis.

A espectroscopia no infravermelho próximo foi eficiente na predição do potencial fisiológico de sementes de lotes de arroz, com 96% de acurácia.

REFERÊNCIAS

- ABUD, H. F.; MESQUISA, C. M. S.; SARMENTO, E. C. S.; MELO, R. S.; LIMA, K. A. P.; SILVA, A. K. F. Análise de imagens de sementes e plântulas de *Vigna radiata* L. **Revista Ciência Agronômica**, v.53, e20207303, 2022.
- ALVARENGA, R.O; MARCOS-FILHO, J.; GOMES-JUNIOR, F. G. Avaliação do vigor de sementes de milho super doce por meio da análise computadorizada de imagens de plântulas. *Revista Brasileira de Sementes*, v. 34, n. 3, p. 488-494, 2012.
- AMANA, H. Z.; JOSHI, R.; MASITHOH, R. E.; CHOUNG, M-G.; KIM, K. H.; CHO, B. K. Nondestructive measurement of anthocyanin in intact soybean seed using Fourier Transform Near-Infrared (FT-NIR) and Fourier Transform Infrared (FT-IR) spectroscopy. **Infrared Physics & Technology**, v. 111, e103477, 2020.
- AMBROSE, A.; KANDPAL, L. M.; KIM, M. S.; LEE, W. H.; CHO, B. K. High speed measurement of corn seed viability using hyperspectral imaging. **Infrared Physics & Technology**, v.75, p.173–179, 2016.
- AMORIM, H. V. Manual de métodos analíticos para o controle da produção de álcool e açúcar. 2. ed. Piracicaba: Esalq-USP, 1996.
- ANTUNES-NETO, A.; MATIAS, F. I.; PASSOS, A. M. A.; ROCHA, N. M. Determinação de sementes de soja esverdeadas por meio de análise de imagens. **Boletim técnico** (Embrapa Milho e Sorgo. Boletim de Pesquisa e Desenvolvimento, 212). 21p, 2020. Disponível em: <<http://www.infoteca.cnptia.embrapa.br/infoteca/handle/doc/1126547>> Acesso em: 25 janeiro 2023.
- ARAÚJO, E. F. L, SANTOS, J. F. dos, SILVA, C. B, da, FILHO, J.M. VIEIRA, R.D. ARAÚJO, J. O.; DIAS, D. C. F. S.; MEDEIROS, A. D.; SILVA, L. J.; NASCIMENTO, W. M. Chickpea seed vigor evaluated by computerized seedling analysis. **Semina: Ciências Agrárias**, v.42, n.1, p.71-86, 2021 (a).
- ARAÚJO, J. D. O., DIAS, D. D. S., DE MEDEIROS, A. D., SILVA, L. J. D., & NASCIMENTO, W. M. Chickpea seed vigor evaluated by computerized seedling analysis. Embrapa Hortaliças-Artigo em periódico indexado (ALICE). 2021
- BAALBAKI, R.; ELIAS, S.; MARCOS-FILHO, J.; McDONALD, M.B. **Seed vigor testing handbook**. Ithaca, NY: AOSA, 2009. 345p. (Contribution 32).
- BADARÓ, A.T., MORIMITSU, F.L., FERREIRA, A.R., CLERICI, M.T.P.S., FERNANDES BARBIN, D., Identification of fiber added to semolina by near infrared (NIR) spectral techniques. *Food Chem.* 289, 195–203. 2019
- BARBOSA, R. M.; VIEIRA, B. G. T. L.; GOMES-JUNIOR, F. G.; VIEIRA, R. D. Image analysis and peanut seeds performance during the production process. *Científica*, v. 44, n. 3, p. 412-420, 2016.
- BARKER, M.; RAYENS, W. Partial mínimos quadrados para discriminação. **Journal of Chemometrics**, v. 17, p. 166-173, 2003.

BAUM, A.; AGGER, J.; MEYER, A. S.; EGOBO, M.; MIKKELSEN, J. D. Rapid near infrared spectroscopy for prediction of enzymatic hydrolysis of corn bran after various pre-treatments. **New Biotechnology**, v.29, p.293-301, 2012.

BAZONI, C. H., IDA, E. I., BARBIN, D. F., & KUROSZAWA, L. E. (2017). Near-infrared spectroscopy as a rapid method for evaluation of physicochemical changes of stored soybeans. *Journal of Stored Products Research*, Reino Unido, v. 73, p. 1-6, 2017.

BERG, S.; KUTRA, D.; KROEGER, T.; STRAELE, B. X.; KAUSLER, B. X.; HAUBOLD, C.; SCHIEGG, M.; ALES, J.; RUDY, M.; EREN, K.; CERVANTES, J. I.; XU, B.; BEUTTENMUELLER, F.; WOLNY, A.; ZHANG, C.; KOETHE, U.; HAMPRECHT, F. A.; KRESHUK, A. Ilastik: Interactive machine learning for (bio)image analysis. *Nature Methods*, v.16, p.1226–1232, 2019.

BERRUETA, L. A., ALONSO-SALCES, R. M., & HÉBERGER, K. Supervised pattern recognition in food analysis. **Journal of chromatography A**, Holanda, v. 1158, n. 1-2, p. 196-214, 2007.

BLANCO, M., VILLARROVA, I., NIR spectroscopy: A rapid-response analytical tool. *TrAC - Trends Anal. Chem.* 21, 240–250. 2002.

BORTOLOTO, R. P.; MENEZES, N. L.; GARCIA, D. C.; MATTIONI, N. M. Teor de proteína e qualidade fisiológica de sementes de arroz. **Bragantia**, v.67, n.2, p.513-520, 2008.

BRASIL. Regras para Análise de Sementes. Brasília: Ministério da Agricultura, Pecuária e Abastecimento, 2009.

BRASIL. Sistema de Consulta à Legislação - SISLEGIS. Ministério da Agricultura, Pecuária e Abastecimento. Secretaria de Defesa Agropecuária. **PORTARIA Nº 111**, set. 2012.

BRERETON, RICHARD G.; LLOYD, GAVIN R. Partial least squares discriminant analysis: taking the magic away. **Journal of Chemometrics**, v. 28, n. 4, p. 213-225, 2014.

BRUNES, A. P. *et al.* Seedling length in wheat determined by image processing using mathematical tools. **Revista Ciencia Agronomica**, v. 47, n. 2, p. 374–379, 2016.

BRUNES, A. P.; ARAÚJO, A. S.; DIAS, L. W.; ANTONIOLLI, G. I. G.; VILLELA, F. A. Rice seeds vigor through image processing of seedlings. *Ciência Rural*, v. 49, n. 8, p. 1-6, 2019.

BURNS, DONALD A., E. W. C. Handbook of Near-Infrared Analysis (D. A. Burns & E. W. Ciurczak (eds.)). CRC Press. 2007.

CASTAN, D. O. C.; GOMES-JUNIOR, F.G.; MARCOS-FILHO, J. Vigor-S, a new system for evaluating the physiological potential of maize seeds. **Scientia Agricola**, v.5, n.2, p.167- 172, 2018.

CHEN, J.; LI, M.; PAN, T.; PANG, L.; YAO, L.; CHANG, J. Rapid and non-destructive analysis for the identification of multi-grain rice seeds with near-infrared spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, v. 219, p. 179-185, 2019.

COHEN, J. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. **Psychological Bulletin**, v.70, n.4, p.213–220, 1968.

COLLINS, T. J. ImageJ for microscopy. **BioTechniques**, v.43, n.1, p.25–30, 2007.

CONAB - Companhia nacional de abastecimento. Acompanhamento da Safra Brasileira, v. 8, n. 9, 2020/21. Disponível em: file:///C:/Users/Usuario/Downloads/E-book_BoletimZdeZSafraZ-Z9oZlevantamento_1.pdf. Acesso em: 16 dic. 2022.

CONAB - Companhia nacional de abastecimento. **Clima adverso influencia na produção de grãos na safra 2022/23, estimada em 310,9 milhões de toneladas. Disponível em : <https://www.conab.gov.br/ultimas-noticias/4885-clima-adverso-influencia-na-producao-de-graos-na-safra-2022-23-estimada-em-310-9-milhoes-de-toneladas>.** Acesso 30 já. 2023.

COZZOLINO, D., Authentication of Cereals and Cereal Products, Advances in Food Authenticity Testing. Elsevier Ltd. 2016.

CUADRADO, J. B. G. Análise de imagem e aprendizagem de máquina na avaliação da qualidade de sementes de arroz. Tese (Mestrado em Fitotecnia) - Universidade Federal de Viçosa, Viçosa, 2020.

DELARMELINO-FERRARESI, L. M.; VILLELA, F. A.; AUMONDE, T. Z. Desempenho fisiológico e composição química de sementes de soja. **Revista Brasileira de Ciências Agrárias**, v.9, n.1, p.14-18, 2014.

DIAS, D.C.F.S.; MARCOS-FILHO, J. Teste de germinação a baixa temperatura. In: In: KRZYZANOWSKI, F.C.; VIEIRA, R. D.; FRANÇA-NETO, J.B.; MARCOS-FILHO, J. **Vigor de sementes: conceitos e testes**. Londrina: ABRATES, p.317-332, 2020.

DIAS, M.A.N., MONDO, V.H.V., CICERO, S.M., N.R. SILVA, C.A.T. da, Vigor tests association as an alternative for precise and efficient assessment of maize seed quality. **Revista Caatinga**, Mossoró, v. 28, n. 3, p. 93 –99, jul. –set., 2015.

DOMINGUES DE SOUZA, L. C., DE SÁ, M. E., DE MORAES, B., MARIA, S., CAMILLO DE CARVALHO, M. A., DA SILVA, M. P., & ABRANTES, F. L. Composição química e nutrientes em sementes das espécies florestais Pente de Macaco, Flor de Paca, Itaúba, Jatobá e Murici Manso. *Bioscience journal*, 478-483. 2012.

EHRHARDT-BROCARD, N. C. M; COELHO, C. M. M. Mobilization of seed storage proteins is crucial to high vigor in common bean seeds. **Ciência Rural**, v.52, n.2, e20200894, 2022.

ERBAŞ, S.; TONGUÇ, M.; SANLI, A. Mobilization of seed reserves during germination and early seedling growth of two sunflower cultivars. **Journal of Applied Botany and Food Quality**, v.89, p.217–222, 2016.

FAN, Y; MA, S.; WU, T. Individual wheat kernels vigor assessment based on NIR spectroscopy coupled with machine learning methodologies. **Infrared Physics & Technology**, v.105, e103213, 2020.

FARHADI, M.; TIGABU, M.; ODÉN, P. Near Infrared Spectroscopy as non-destructive method for sorting viable, petrified and empty seeds of *Larix sibirica*. *Silva Fennica*, v.49, n.5, 2015.

FERREIRA, O. J. M.; ROCHA, L. A. S.; SILVA-MANN, R.; TORRES, M. F. O.; SOUZA, J. L.; DANTAS, S. J.; SANTO, R. C.; SANTOS, J. P. F. Tecnologia de análise de imagens para a seleção de sementes crioulas de milho. **Global Science and Technology**, v.13, n.02, p.28-38. 2020.

FOLHA, THAIS OLIVEIRA. Uso da espectroscopia no infravermelho próximo (NIR) para a avaliação de parâmetros de qualidade de farinha de mandioca (Master's thesis, Universidade Federal de Pernambuco). 2014.

GAITÁN-JURADO, A. J., GARCÍA-MOLINA, M., PEÑA-RODRÍGUEZ, F., & ORTIZSOMOVILLA, V. Near infrared applications in the quality control of seed cotton. *Journal of Near Infrared Spectroscopy*, Reino Unido, v. 16, n. 4, p. 421-429, 2008.

GOMES-JUNIOR, F. G.; CHAMMA, H. M.C. P.; CICERO, S. M.; Automated image analysis of seedlings for vigor evaluation of common bean seeds. *Acta Scientiarum. Agronomy*, v. 36, n. 2, p. 195-200, 2014.

GOMES-JUNIOR; F. G. Análise computadorizada de imagens de plântulas. In: KRZYZANOWSKI, F. C.; VIEIRA, R. D.; FRANÇA-NETO, J. B.; MARCOS-FILHO, J. (ed.). Vigor de sementes: conceitos e testes. Londrina: ABRATES, p.139-181, 2020.

GRZYBOWSKI, C. R. DE S.; VIEIRA, R. D.; PANOBIANCO, M. Testes de estresse na avaliação do vigor de sementes de milho. **Revista Ciencia Agronomica**, v. 46, n. 3, p. 590–596, 2015.

GU, J.; CHAO, H.; GAN, L.; GUO, L.; ZHANG, K.; LI, Y.; LI, M. Proteomic dissection of seed germination and seedling establishment in *Brassica napus*. **Frontiers in Plant Science**, v.7, p.1–19, 2016.

GUIMARÃES, G. C. Espectroscopia no infravermelho próximo para classificação de sementes de café quanto à qualidade, origem e cultivar. Tese (doutorado) – Universidade Federal de Lavras, Lavras. 2016.

HENNING, F. A.; MERTZ, L. M.; JACOB-JUNIOR, E. A.; MACHADO, R. D.; FISS, G.; ZIMMER, P. D. Composição química e mobilização de reservas em sementes de soja de alto e baixo vigor. **Bragantia**, v.69, n.3, p.727-734, 2010.

HOURLANT, P.; BAETEN, V.; MORALES, M.T.; MEURENS, M.; APARICIO, R. Oil and fat classification by selected bands of near-infrared spectroscopy. **Applied Spectroscopy**, v.54, p.1168-1174, 2000.

HUANG, Z., SHA, S., RONG, Z., CHEN, J., HE, Q., KHAN, D. M., ZHU, S. Feasibility study of near infrared spectroscopy with variable selection for non-destructive determination of quality parameters in shell-intact cottonseed. *Industrial crops and products*, Holanda, v. 43, p. 654-660, 2013.

JOLLIFFE, I. T.; CADIMA, J. Principal component analysis: a review and recent developments. **Philosophical Transactions of the Royal Society A**, v. 374, n. 20150202, p. 1-16, 2016.

KANDPAL, L. M.; LOHUMI, S.; KIM, M. S.; KANG, J. S.; CHO, B.-K. Near-infrared hyperspectral imaging system coupled with multivariate methods to predict viability and vigor in muskmelon seeds. *Sensors and Actuators B: Chemical*, v.229, p.534–544, 2016.

KRZYZANOWSKI, F.C.; FRANÇA-NETO, J.B.; GOMES-JUNIOR, F.G.; NAKAGAWA, J. Testes de vigor baseados em desempenho de plântulas. **Vigor de sementes: conceitos e testes**. Londrina: ABRATES, p.79-140, 2020.

KUMAR, N.; BANSAL, A.; SARMA, G. S.; RAWAL, R. K. Chemometrics tools used in analytical chemistry: **An overview**. v. 123, p. 186-199, 2014.

KUSUMANINGRUM, D., LEE, H., LOHUMI, S., MO, C., KIM, M. S., CHOA, B. Non-destructive technique for determining the viability of soybean (*Glycine max*) seeds using NIR spectroscopy. **Journal of the Science of Food and Agriculture**, v. 98, p. 1734-1742, 2018.

LARIOS, G.; NICOLODELLI, G.; RIBEIRO, M.; CANASSA, T.; RODRIGUES DOS REIS, A.; OLIVEIRA, S.; ALVES, C. Z.; MARANGONI, B. S.; CENA, C. Soybean seed vigor discrimination by infrared spectroscopy and machine learning algorithms. *Analytical Methods*, 12, p. 4303-4309, 2020.

LEE, J. H., & CHOUNG, M. G. Nondestructive determination of herbicide-resistant genetically modified soybean seeds using near-infrared reflectance spectroscopy. *Food chemistry*, 126(1), 368-373. 2011

LIMÃO, M.A.R. Testes de vigor, atividade enzimática e espectroscopia no infravermelho próximo em sementes de lentilha (*Lens culinaris* Medik.). **Dissertação** (Mestrado em Fitotecnia) - Universidade Federal de Viçosa, Viçosa, 2022.

LORENTZ, L. H., & NUNES, U. R. Relações entre medidas de qualidade de lotes de sementes de arroz. *Revista Ciência Agronômica*, 44, 798-804. 2013.

MAGUIRE, J.D. Speed of germination-and in selection and evaluation for seeding emergence and vigor. **Crop Science**, v.2, n.2, p.176-177, 1962.

MARCOS-FILHO, J. Seed vigor testing: an overview of the past, present and future perspective. *Scientia Agricola*, v.72, n.4, p.363-374, 2015.

MARCOS-FILHO, J. **Fisiologia de sementes de plantas cultivadas**. Londrina: ABRATES, 659p, 2015 (a).

MARCOS-FILHO, J. System computerized analysis of seedling images (SVIS®) for evaluation of seed vigor. *Revista Brasileira de Sementes*, v. 20, p. 40-44, 2010.

MARCOS-FILHO, J. Teste de envelhecimento acelerado. In: KRZYZANOWSKI, F.C.; VIEIRA, R. D.; FRANÇA-NETO, J.B.; MARCOS-FILHO, J. **Vigor de sementes: conceitos e testes**. Londrina: ABRATES, p.182-244, 2020.

MARCOS-FILHO, J.; KIKUTI, A. L. P.; LIMA, L. B. Métodos para avaliação do vigor de sementes de soja, incluindo análise computadorizada de imagens. **Revista Brasileira de Sementes**, v.31, n.1, p.102-112, 2009.

MARCOS-FILHO, J.; KRZYZANOWSKI, F.C.; VIEIRA, R.D.; FRANÇA NETO, J. B. Vigor de sementes: conceitos e testes. **ABRATES**. [S.l: s.n.], p. 2–3, 1999.

MATSUO, M. et al. Practical application of near-infrared spectroscopy for determining rice amylose content at grain elevator. *Azojete*, v. 14, p. 95-100, 2018.

MAYRINCK, L. G. Uso da espectrometria no infravermelho próximo na avaliação da qualidade fisiológica de sementes de algodão. Dissertação (mestrado) – Universidade Federal de Lavras, Lavras. 2018.

MCCORMAC, A. C. ; KEEFE, P. D. ; DRAPER, S. R. Automated vigour testing of field vegetables using image analysis. **Seed Science and Technology**, v. 18, n. 1, p. 103–112, 1990.

MEDEIROS, A.D.DE *et al.* Interactive machine learning for soybean seed and seedling quality classification. **Scientia Agricola**, p. 1–13, 2020.

MEDEIROS, A. D. de; PEREIRA, M. D. SAPL®: a free software for determining the physiological potential in soybean seeds. *Pesquisa Agropecuária Tropical*, v. 48, n. 3, p. 222– 228, 2018.

MEDEIROS, A. D. *et al.* Vigor of maize seeds determined by a free image analysis system. **Revista Ciencia Agronomica**, v. 50, n. 4, p. 616–624, 2019.

MEDEIROS, A. D., PEREIRA, M.D, SILVA, J.A. Processamento digital de imagens na determinação do vigor de sementes de milho. **Revista Brasileira de Ciências Agrárias** ISSN (online) 1981-0997 v.13, n.3, e5540, 2018.

MEDEIROS, A. D.; CAPOBIANGO, N.P.; SILVA, J. M.; SILVA, L. J.; SILVA, C. B.; DIAS, D. C. F. S. Interactive machine learning for soybean seed and seedling quality classification. **Scientific Reports**, v.10, n.1, p.11267, 2020 (a).

MEDEIROS, A. D.; PINHEIRO, D. T.; XAVIER, W. A.; SILVA, L. J.; DIAS, D. C. F. S. Classificação de qualidade de sementes de pinhão manso utilizando imagens radiográficas e aprendizado de máquina. *Industrial Crops and Production*, v.146, e112162, 2020 (b).

MEDEIROS, M. L, S. Application of Near Infrared spectroscopy and hyperspectral imaging to quantify the oil content and classify Brassica seeds (Doctoral dissertation, Universidade Estadual de Campinas (UNICAMP). Faculdade de Engenharia de Alimentos). 2021.

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

NAKAGAWA, J. **Testes de vigor baseados no desempenho das plântulas**. In: KRZYZANOWSKI, F.C.; VIEIRA, R.D.; FRANÇA NETO, J.B. (Ed.). Vigor de sementes: conceitos e testes. Londrina: ABRATES. p. 2.1-2.24. 1999.

NIE, P.; ZHANG, J.; FENG, X.; YU, C.; HE, Y. Classification of hybrid seeds using near- 83 infrared hyperspectral imaging technology combined with deep learning. *Sensors and Actuators B: Chemical*, v.296, e.126630, 2019.

NOLASCO, I.M.P. Caracterização e Identificação de carne de aves por métodos não destrutivos. Universidade Estadual de Campinas. 2019.

PASQUALLI, LL. **Qualidade de sementes de arroz irrigado submetidas a diferentes temperaturas na secagem estacionaria.** 37f. Dissertação (Mestrado em Agronomia) – Universidade Federal de Santa Maria, 2005.

PASQUINI, C. Near infrared spectroscopy: A mature analytical technique with new perspectives—A review. *Analytica. Chimica Acta*, v.1026, p.8–36, 2018.

PEREIRA, MÁRCIO DIAS; REIS, JOÃO ANTÔNIO DOS; FERRARI, C. DOS S. Processamento digital de imagens de plântulas na avaliação do vigor de sementes de *Moringa oleifera* Lam. *Revista Ciencia Agronomica*, v. 30, n. 2, p. 291–306, 2020.

PESKE, S.T.; ROSENTHAL, M.A.; ROTA, G.R.M. Sementes: fundamentos científicos e tecnológicos. Pelotas: Universidade Federal de Pelotas. Editora Universitária, 2003.

QIU Z, CHEN J, ZHAO Y, ZHU S, HE Y, ZHANG C. Variety Identification of Single Rice Seed Using Hyperspectral Imaging Combined with Convolutional Neural Network. *Applied Sciences*. 8(2):212. 2018 (a).

QIU, G.; LÜ, E.; LU, H.; XU, S.; ZENG, F.; SHUI, Q. Single-kernel FT-NIR spectroscopy for detecting supersweet corn (*Zea mays* L. Saccharata sturt) seed viability with multivariate data analysis. *Sensors*, v.18, n.4, e.1010, 2018 (b).

QUEIRÓZ, T.N.; VALIGUZSKI, A. L.; BRAGA, C. DOS S.; SOUZA, S. A. M.; DA ROCHA, A. M. Avaliação da qualidade fisiológica de sementes de variedades tradicionais de milho. *Revista da Universidade Vale do Rio Verde*, v.17, n.1, 2019.

R CORE TEAM. R Development Core TeamR: A Language and Environment for Statistical Computing, 2020.

RAVAZZI, C.G. Identificação e quantificação de adulterantes de Whey protein concentrado empregando espectroscopia no infravermelho próximo e resolução multivariada de curvas, in: Dissertação de Mestrado (Química Analítica). Universidade Estadual de Campinas. p. 75. 2019

REDDY, P.; GUTHRIDGE, K.M.; PANOZZO, J.; LUDLOW, E.J.; SPANGENBERG, G.C.; ROCHFORD, S.J. Near-Infrared Hyperspectral Imaging Pipelines for Pasture Seed Quality Evaluation: An Overview. *Sensors*, 22 (5), 2022.

REIS, J. A. V. D. Sistema de análise de plântulas (SAPL) na determinação do vigor de sementes de *Moringa oleifera* Lam (Master's thesis, Brasil). 2017

RIBEIRO, J. P. O.; MEDEIROS, A. D.; CALIARI, I. P.; TRANCOSO, A. C. R.; MIRANDA, R. M.; FREITAS, F. C. L.; SILVA, L. J.; DIAS, D. C. F. dos S. FT-NIR and linear discriminant analysis to classify chickpea seeds produced with harvest aid chemicals. *Food Chemistry*, v.341, e.128324, 2020.

RINNAN, Å.; BERG, F. VAN DEN; ENGELSEN, S. B. Review of the most common pre-processing techniques for near-infrared spectra. **TrAC Trends in Analytical Chemistry**, v. 28, n. 10, p. 1201–1222, nov. 2009.

ROCHA, CRM da et al. Avaliação do vigor de sementes de girassol por meio de análise de imagens de plântulas. **Ciência Rural**, v. 45, n. 6, pág. 970-976, 2015.

SAKO, Y.; McDONALD, M. B.; FUJIMURA, K.; EVANS, A. F.; BENNETT, M. A. A system for automated seed vigour assessment. **Seed Science and Technology**, v.29, p.625-636, 2001.

SANTOS, M. D. F. D. P. Desenvolvimento e Validação de Métodos de Espectroscopia no Infravermelho Próximo e Médio para Caracterização de Lamas de ETAR para Uso Agrícola (Doctoral dissertation, Dissertação de Mestrado, Universidade Técnica de Lisboa, Portugal). 2007

SAVITZKY, A.; GOLAY, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. **Analytical Chemistry**, v. 36, n. 8, p. 1627–1639, 1964.

SENA, D.V.A.; ALVES, E.U.; MEDEIROS, D.S. Vigor de sementes de milho cv. 'Sertanejo' por testes baseados no desempenho de plântulas. **Ciência Rural**, v. 45, n. 11, p. 1910-1916, 2015.

SHENK, J. S.; WESTERHAUS, M. O.; WORKMAN, J. J. 2007. Application of NIR Spectroscopy to Agricultural Products. In: D. A. Burns and E. W. Ciurczak (eds.) Handbook of Near-Infrared Analysis, Third Edition. Practical Spectroscopy No. 35. p 347-386. CRC Press.

SHRESTHA, S., KNAPIČ, M., ŽIBRAT, U., DELEURAN, L. C., GISLUM, R. Single seed near-infrared hyperspectral imaging in determining tomato (*Solanum lycopersicum* L.) seed quality in association with multivariate data analysis. **Sensors and Actuators B: Chemical**, Holanda, v. 237, p. 1027-1034, 2016.

SILVA, K.C.N. Avaliação do vigor de sementes crioulas de milho (*Zea Mays* L.) utilizando testes tradicionais e análise de imagens. **Dissertação (Mestrado em Ambiente, Tecnologia e Sociedade)**, Mossoró RN, 2020.

SILVA, L. J.; MEDEIROS, A. D.; OLIVEIRA, A. M. S. SeedCalc, a new automated R software tool for germination and seedling length data processing. **Journal of Seed Science**, v.41, n.2, p.250-257, 2019.

SILVA, PP da et al. Avaliação do vigor de sementes de abóbora por meio de análise computadorizada de imagens. **Journal of Seed Science**, v. 39, n. 2, pág. 159-165, 2017.

SOARES, S. F. C.; MEDEIROS, E. P.; PASQUINI, C.; DELELIS MORELLO, C.; GALVÃO, R. K. H.; ARAÚJO, M. C. U. Classification of individual cotton seeds with respect to variety using near-infrared hyperspectral imaging. **Analytical Methods**, v. 8, n. 48, p. 8498-8505, 2016.

SOMMER, C.; STRAEHLE, C.; KOTHE, U.; HAMPRECHT, F. A. Ilastik: Interactive learning and segmentation toolkit. **International Symposium on Biomedical Imaging: From Nano to Macro**. v.978, n.1, p.4244, 2011.

SOSBAI. Sociedade Sul-Brasileira de Arroz Irrigado. Arroz irrigado: Recomendações Técnicas da Pesquisa para o Sul do Brasil. 205p. 2018.

VENKATESAN, S.; MASILAMANI, P.; JANAKI, P.; EEVERA, T.; SUNDARESWARAN, S.; RAJKUMAR, P. Role of Near - Infrared Spectroscopy in Seed Quality Evaluation: A Review. **Agricultural Reviews**, R1960, p. 1-10, 2020.

VISSCHERS, I. G. S.; VAN D. A. M.; PETERS, J. L. Um método objetivo de triagem de alto rendimento para quantificação de danos por tripes usando Ilastik e ImageJ. **Entomologia Experimentalis Applicata**. v.166, p.508-515, 2018.

VREŠAK, M.; HALKJAER OLESEN, M.; GISLUM, R.; BAVEC, F.; RAVN JØRGENSEN, J. The use of image-spectroscopy technology as a diagnostic method for seed health testing and variety identification. **PLOS ONE**, v.11, n.3, e0152011, 2016.

WANG, Q.; LIU, H.; SHI, A.; HU, H.; LIU, L.; WANG, L.; YU, H. Review on the processing characteristics of cereals and oilseeds and their processing suitability evaluation technology. **Journal of Integrative Agriculture**, v.16, n.12, p.2886–2897, 2017.

WANG, Y.; PENG, Y.; ZHUANG, Q.; ZHAO, X. Feasibility analysis of NIR for detecting sweet corn seeds vigor. **Journal of Cereal Science**, v.93, e102977, 2020.

WORKMAN, J. JR.; WEYER, L. **Practical Guide to Interpretive Near-Infrared Spectroscopy**. CRC Press, Taylor & Francis Group, New York, NY. 346p. 2007.

WORMA, M.; SEGATTO, C.; STEFEN, C.; BUBA, G. P.; LEOLATO, L. S. Qualidade fisiológica de sementes de milho produzidas com adubação biológica e bioestimulante em diferentes preparos de solo. **Revista Engenharia na Agricultura - Reveng**, v.27, n.3, p.187-194, 2019.

WU, D.; CHEN, X.; SHI, P.; WANG, S.; FENG, F.; HE, Y. Determination of a linolenic acid and linoleic acid in edible oils using near-infrared spectroscopy improved by wavelet transform and uninformative variable elimination. **Analytica Chimica Acta**, v.634, p.166-171, 2009.

XIA, Y.; XU, Y.; LI, J.; ZHANG, C.; FAN, S. Recent advances in emerging techniques for non-destructive detection of seed viability: A review. **Artificial Intelligence in Agriculture**, v.1, p.35–47, 2019.

XU, R.; HU, W.; ZHOU, Y.; ZHANG, X.; XU, S.; GUO, Q.; QI, P.; CHEN, L.; YANG, X.; ZHANG, F.; LIU, L.; QIU, L.; WANG, J. Use of near-infrared spectroscopy for the rapid evaluation of soybean [*Glycine max* (L.) Merri.] water soluble protein content. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, v. 224, n. 117400, p. 1-8, 2020.

YANG, X.; HONG, H.; YOU, Z.; CHENG, F. Spectral and image integrated analysis of hyperspectral data for waxy corn seed variety classification. **Sensors**, v.15, n.7, p.15578–94, 2015.

YATES, S. C.; GROENEBOOM, N. E.; COELLO, C.; LICHTENTHALER, S.; KUHN, P. H.; DEMUTH, H. U.; RUBSAMEN, M. H.; ROBNER, S.; LEERGAARD, T.; KRESHUK, A.; PUCHADES, M. A.; BJAALIE, J. G. QUINT: Fluxo de trabalho para quantificação e análise espacial de características em imagens histológicas do cérebro de roedores. **Frontier Neuroinformatic**, v.13, p.1-14, 2019.

ZHANG, J.; LI, M.; PAN, T.; YAO, L.; CHEN, J. Purity analysis of multi-grain rice seeds with non-destructive visible and near-infrared spectroscopy. *Computers and Electronics in Agriculture*, v.164, e104882, 2019.

ZHAO, Y.; ZHANG, C.; ZHU, S.; GAO, P.; FENG, L.; HE, Y. Non-destructive and rapid variety discrimination and visualization of single grape seed using near-infrared hyperspectral imaging technique and multivariate analysis. **Molecules**, v.23, n.6, p.1352, 2018.

ZHU, S.; ZHANG, J.; CHAO, M.; XU, X.; SONG, P.; ZHANG, J.; HUANG, Z. A rapid and highly efficient method for the identification of soybean seed varieties: hyperspectral images combined with transfer learning. *Molecules*, v.25, n.1, p.152, 2019.

ZONTA, J., ABREU, G., SANTIAGO, C., NOGUEIRA, L., FONTES, M., de MELO, P. A. F. R., & RODRIGUES, A. Qualidade de sementes de arroz de produção e uso próprios por agricultores familiares no Estado do Maranhão. 2022.