ARTICLE

Near infrared spectroscopy and seedling image analysis to evaluate the physiological potential of *Urochloa decumbens* (Stapf) R.D. Webster seeds

Lucas Rodrigues de Souza[1] iD, **Marcelo Augusto Rocha Limão**[1] iD, **Daniel Teixeira Pinheiro**[1*] iD, **Gabriel Cordeiro de Oliveira Peris**[1] iD, **Denise Cunha Fernandes dos Santos Dias**[1] iD, **Luiz Antônio dos Santos Dias**[1] iD

ABSTRACT: The demand for techniques that make it possible to assess the physiological potential of seeds quickly and assertively make near-infrared spectroscopy (FT-NIR) and seedling analysis using ILASTIK software promising tools. The aim of this study was to evaluate the physiological potential of *Urochloa decumbens* seeds using near-infrared spectroscopy (FT-NIR) and ILASTIK software. Seeds from 10 lots of *U. decumbens* were classified according to their physiological potential (germination and vigor) and then FT-NIR spectra were obtained from individual seed samples, the original spectra being pre-processed with different dispersion correction methods used for the construction of a classification model through partial least squares discriminant analysis (PSL-DA). For the ILASTIK evaluation, the seedlings were photographed at 7 and 14 days of germination and the trained classifier was applied to the images, generating data on the numbers of strong seedlings, weak seedlings and non-germinated seeds. With data from the FT-NIR technique pre-processed by the 2nd derivative of Savitzky-Golay, it was possible to obtain a classification model with high efficiency to discriminate the classes regarding the physiological potential of the seeds. ILASTIK was efficient to classify seeds according to their physiological potential after only 7 days of germination. FT-NIR and ILASTIK analyses are non-destructive and fast alternatives, with great potential for quality control of *U. decumbens* seed lots.

Index terms: brachiaria, ILASTIK, FT-NIR technique, physiological tests.

RESUMO: A demanda por técnicas que permitam avaliar o potencial fisiológico das sementes de forma rápida e assertiva fazem da espectroscopia no infravermelho próximo (FT-NIR) e da análise de plântulas pelo *software* ILASTIK ferramentas promissoras. O objetivo do trabalho foi avaliar o potencial fisiológico de sementes de *Urochloa decumbens* por meio da espectroscopia no infravermelho próximo (FT-NIR) e do software ILASTIK. Sementes de 10 lotes de *U. decumbens* foram classificadas quanto o potencial fisiológico (germinação e vigor) e depois foram obtidos espectros FT-NIR de amostras individuais de sementes, sendo os espectros originais pré-processados utilizando diferentes métodos de correção de dispersão utilizados para a construção de um modelo de classificação por meio da PLS-DA. Para a avaliação do ILASTIK, as plântulas foram fotografadas aos 7 e 14 dias de germinação e o classificador treinado foi aplicado às imagens, gerando dados de número de plântulas fortes, plântulas fracas e sementes não germinadas. Com os dados da técnica FT-NIR pré-processados pela 2ª derivada de Savitzky-Golay, foi possível obter um modelo de classificação com alta eficiência para discriminar as classes quanto ao potencial fisiológico das sementes. O ILASTIK foi eficiente para a classificação das sementes quanto ao potencial fisiológico com apenas 7 dias de germinação. As análises pelo FT-NIR e ILASTIK são alternativas não destrutivas e rápidas, com grande potencial para o controle de qualidade de lotes de sementes de *U. decumbens*.

Termos para indexação: braquiária, ILASTIK, técnica FT-NIR; testes fisiológicos.

INTRODUCTION

In Brazil, plants of the genus *Urochloa* were introduced with the aim of implementing pastures, since they have high compatibility with tropical soils and climates (Paula et al., 2017). In this context, about 80% of the pastures in the country are occupied by plants of this genus, which has become a protagonist and a strong ally in the country's agricultural production (Ferreira et al., 2021; EMBRAPA, 2022). Among the species, *Urochloa decumbens* has been showing efficiency in weed suppression in areas cultivated with perennial crops (Martinelli et al., 2018) and in helping control nematodes (Asmus and Cruz, 2020), among other advantages.

U. decumbens seeds have physiological dormancy in the post-harvest, which can impair the emergence of seedlings in the field, leading to uneven stand (Batista et al., 2015). In addition, there is uneven maturation of the seeds, which affects both the physical and the physiological quality. The physiological quality of *U. decumbens* seeds is mainly evaluated by the germination test, which provides the results after 21 days, and by the tetrazolium test, which is laborious and whose interpretation is subjective, requiring training and experience on the part of the analyst. Currently, there is a demand for faster and simpler methods that optimize decision making regarding the approval or discard of the lots produced in each season. Although these new methodologies do not replace the traditional tests used before commercialization, simple, fast and high-yield methods are fundamental for a rapid pre-selection in the quality control of lots.

In this context, the evaluation of the vigor of seed lots based on seedling performance through image analysis software is a promising tool, mainly due to the agility in obtaining the results and accuracy of the obtained data (Gomes-Junior, 2020). Among the various software programs available for image analysis of seeds and seedlings, ILASTIK has open source and a dynamic interface, allowing the user to manually define and classify the image resources, defining the objective autonomously and with easy handling (Berg et al., 2019). Furthermore, it allows data manipulation based on interactive learning from the acquisition of images, which depends only on a digital camera (Berg et al., 2019). Medeiros et al. (2020) observed high accuracy of ILASTIK to classify soybean seeds and seedlings as to appearance and physiological potential, quickly and non-destructively.

Another interesting tool that has been tested for rapid prediction of the physiological potential of seeds is near-infrared spectroscopy (NIR), which has the advantages of being simple, non-destructive and non-subjective (Ribeiro et al., 2023). In general, this technique is based on the absorption of wavelengths of electromagnetic radiation with the range of 780-2,500 nm, enabling the formation of spectra that correlate with the functional groups of the molecules (C-H, N-H and O-H) present in the sample (Xia et al., 2019). For the evaluation of seed lots, the use of NIR allows the construction of graphs from the obtained spectra, which contain peaks of wavelengths referring to the chemical compounds that are altered according to the level of quality of the seed. Thus, this technique has been studied and proved to be viable for predicting the physiological potential of seeds of various species such as chickpea (Ribeiro et al., 2021), corn (Andrade et al., 2020; Andriazzi et al., 2023), *Brassica* spp. (Medeiros et al., 2022), among others.

Considering that fast, non-invasive methods with less human interference in the evaluation of seed quality are currently highly desirable and demanded by the seed industry for optimizing decision-making regarding the approval or discard of lots, the present study aimed to evaluate the physiological potential of *U. decumbens* seeds by means of near-infrared spectroscopy (FT-NIR) and ILASTIK software.

MATERIAL AND METHODS

The study was carried out at the Seed Research Laboratory of the Agronomy Department of *Universidade Federal de Viçosa*, in Viçosa, Minas Gerais, Brazil. Ten lots of *Urochloa decumbens* seeds, provided by the *Empresa Brasileira de Pesquisa Agropecuária* (Embrapa) and produced in the 2021/2022 harvest, were used. The study was divided into three assays:

Assay I - Characterization of the physiological potential of U. decumbens seeds

Moisture content: determined with two replications of 20 seeds by the oven method at 105 ± 3 °C for 24 hours. The results were expressed as percentage (wet basis) (Brasil, 2009).

Germination test: conducted with four replications of 25 seeds, which were later used for the FT-NIR analysis. The seeds were distributed in gerbox boxes with germination paper moistened with an amount of KNO_3 equivalent to 2.5 times the weight of the dry paper. The gerbox boxes were kept in a germinator at alternating temperatures of 20-35 °C and the evaluations were carried out daily until the twenty-first day after setting up the test (Brasil, 2009). Germination was computed by the percentage of root protrusion (greater than 2 mm) and normal seedlings obtained at 21 days.

First germination count: evaluated together with the germination test, by determining the percentage of normal seedlings on the seventh day (Brasil, 2009).

Germination indices: data obtained in the daily counts performed in the germination test were used to calculate the germination speed index (GSI) (Maguire, 1962) and time to germination of 50% of the lot (t50). Both indices were calculated using the SeedCalc package of R software (Silva et al., 2019).

Seedling dry matter: four replications of 20 seedlings obtained at 14 days per lot were dried in a forced air circulation oven at 72 °C until weight stabilization, and the results were expressed in g.seedling^{-1}.

Experimental design and statistical analysis: The experimental design was completely randomized (CRD) with 10 lots and four replications. Data were subjected to analysis of variance (ANOVA). After confirming the normal distribution of errors by the Shapiro-Wilk test, the means were compared by the Scott-Knott test at 5% probability level. All analyses were performed in R statistical software (R Core Team, 2022).

Assay II - Near-infrared spectroscopy - NIR and its relationship with the physiological potential of seeds

Acquisition of FT-NIR spectra: 100 seeds per lot were used for this analysis and later used for the germination test, totaling 1,000 seeds analyzed individually. The seeds were deposited in a circular support with size and shape suitable for *U. decumbens* seeds. The spectral data of each seed were obtained using a Fourier-transform spectrometer – FT-NIR (Thermo Scientific Antaris II). The reflectance spectra were expressed as log (1/R), with R being the reflectance. For each seed, 3,111 points per spectrum were collected, within the wavelength range from 1,000 to 2,500 nm. 200 successive scans of the spectra of each lot were performed, in 30 seconds for each seed.

Pre-processing algorithms: the original spectral data were pre-processed with the methods Standard Normal Variate (SNV), Multiplicative Scatter Correction (MSC), and 1st and 2nd derivatives of Savitzky-Golay (SG), using the window of 11 variables in wavelengths, which aims to reduce the number of scores.

Development of models for classification: U. decumbens seed lots were classified according to the result of the germination test and the percentage of vigorous seedlings. Three classes were defined: high physiological potential (HPP) (germination ≥ 69%), medium physiological potential (MPP) (53% ≤ germination ≤ 68%) and low physiological potential (LPP) (germination ≤ 52%). Thus, the number of lots per class was composed of 200 spectra for the HPP class, 400 for the MPP class and 400 for the LPP class.

Model validation: The models were constructed using 70% of the data for training and the remaining 30% for validation. The pre-processing models with highest accuracy were used to predict the performance of the lots in the validation set.

Experimental design: The germination test after FT-NIR analysis was performed as described in Assay I.

The classes were assigned to the NIR spectra, according to the physiological performance of the lots established by the vigor tests. An exploratory analysis was performed with the original spectra and the mean of the classes, followed by the PCA. For each pre-processing, a classification model was generated using partial least squares discriminant analysis (PLS-DA) (Barker and Rayens, 2003). The performance of the models was evaluated by accuracy and Kappa in both tests. In addition, the wavelength ranges that were most important for each vigor class in the construction of the classification model were identified. PLS-DA was performed using the packages caret, prospectr, viridis, patchwork, factoMineR and ggplot2 in R software (R Core Team, 2022).

Assay III - Computerized analysis of seedlings by ILASTIK software

It was performed according to the seedling length test, with eight replications of 10 seeds from each lot. The seeds were distributed in the upper third of the germination paper, in the longitudinal direction. They were placed to germinate as described for the germination test (Assay I). Subsequently, seedlings at 7 and 14 days were transferred to a photographic base made of blue satin vinyl foam (E.V.A) sheet. Images were acquired through photographs, using a Nikon Coolpix P510 digital camera, configured in 16 Megapixels. The camera was kept at a height of 40 cm and angle of 90° in relation to the photographic base, using a copystand support.

The images obtained were processed by ILASTIK software, which is connected to a CellProfiler module, free and open access (https://www.ilastik.org/download), allowing interactive categorization of bioimages. In the segmentation, the color and pixel classification tool were used, initially determining two classes of segmentation: "seed or seedling" (region of interest) and "background" (discard region). To create the probability maps, pixels belonging to the regions of each segmented class were trained to identify the seedlings according to the colors defined in the segmentation. Then, the software was trained once again to recognize three classes of seedlings: strong seedlings (well-developed root and shoot, greater than 5 cm); weak seedlings (absence, underdevelopment or deformation of some essential structure); and non-germinated seeds (Figure 1). The trained classifier was applied to all images and the probability maps were exported, generating data on the numbers of strong and weak seedlings and non-germinated seeds.

Figure 2 illustrates the training process of ILASTIK software, showing initially the images of the seedlings of the germination test on the blue background (I). Next, rendering was applied for segmentation and the region of interest (ROI) was improved to acquire the probability map (II). After rendering the ROI, the software identifies the individual seedlings and seeds with the probability map (III), and finally the classification of each seed and seedling is predicted from the training colors of the respective classifier groups (IV).

Figure 1. Pattern of strong seedlings and weak seedlings used for classification by ILASTIK software.
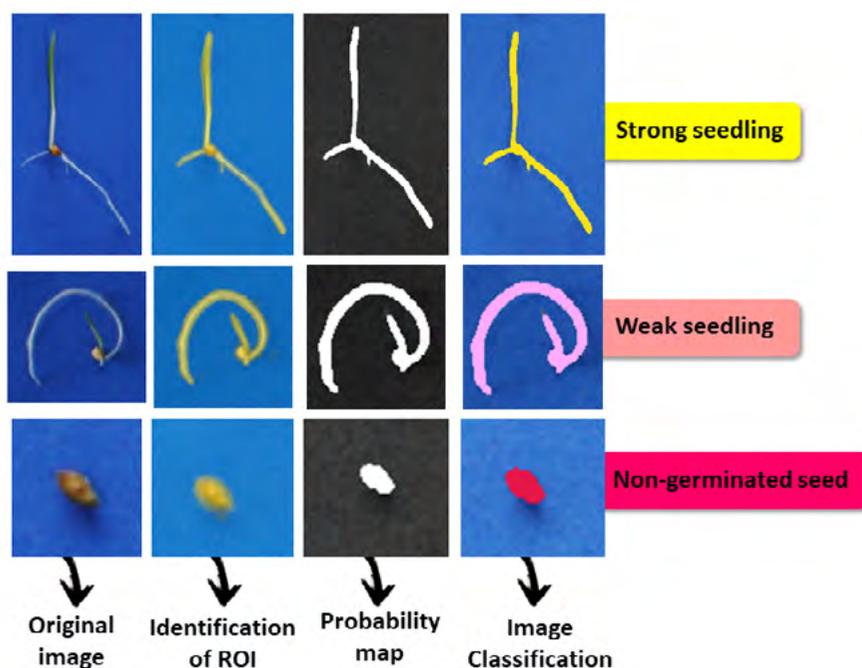
Figure 2. Representative scheme of the steps of interactive machine learning and classification of physiological quality by ILASTIK software of the different lots of *U. decumbens* seeds.

Subsequently, multivariate principal component analysis (PCA) was performed for the attributes generated by ILASTIK and for the physiological tests evaluated in Assay I. An "n x p" matrix was obtained, where "n" corresponds to the number of treatments (n = 10) and "p" corresponds to the number of variables analyzed (p = 12). Eigenvalues and eigenvectors were calculated from the covariance matrices and plotted in two-dimensional graphs (category ordering diagram and correlation circle). The statistical software R was used for this analysis (R Core Team, 2022).

RESULTS AND DISCUSSION

Physiological characterization of the lots highlighted the uniformity of seed moisture, which remained between 10 and 13% in all lots evaluated. This pattern is important because moisture levels with high variation interfere with the germination capacity of the lots, which can alter results (Marcos-Filho, 2015a) (Table 1).

Through the physiological tests, it was possible to observe, based on both root protrusion (RP) and percentage of normal seedlings (NS) obtained by the germination test at 21 days, that lots 2 and 8 were superior to the other lots evaluated and classified as of high physiological potential (HPP). Tests related to germination speed, such as first germination count (FGC), germination speed index (GSI) and t50, allowed confirming, in general, higher physiological potential of lots 2 and 8 and lower physiological potential of lots 3, 6 and 10. It is important to emphasize that, in addition to the germination potential, seeds must show a high speed of emergence in the field, which contributes for them to have less exposure to adverse factors such as temperature, pest attacks, water deficit, among others (Marcos-Filho, 2015b). These results are complemented by those observed for seedling dry matter (SDM), as lots 2 and 8 showed higher means and lots 3, 6 and 10, lower means (Table 1). Therefore, it can be considered that lots 2 and 8 and lots 3, 6 and 10 have higher and lower physiological potentials when compared to the other intermediate lots analyzed, respectively. Lots 1, 4, 5, 7 and 9 showed intermediate physiological potential (Table 1).

Figure 3A shows the total of 1,000 original NIR spectra that correspond to the 10 lots of *U. decumbens*. The means of the original data are presented in Figure 3B, where it is possible to observe differences in three levels of

physiological potential of the seeds. Principal component analysis (PCA) was carried out with the 1,000 spectra, with the objective of performing exploratory data analysis (Figure 4). The central ordering diagram shows the scores and arrangement of the spectra, differentiated by the respective classes of physiological potential (Figure 4A). The analysis of the two principal components (PC1 and PC2) explained 99.8% of the total variation of the data of physiological potential of *U. decumbens* seeds, confirming the efficiency of the technique for separating the different levels of physiological potential of the seeds.

Chemometric transformation techniques used in spectral data have been very efficient to reduce the noise of the samples and equipment. Pre-processing of the data before classification enables the refinement of the spectra, to subsequently obtain the best fit and training model (Rinnan et al., 2009). Different data processing methods were tested, such as multiplicative scatter correction (MSC), standard normal variate (SNV) and derivative methods (Savitzky-Golay), which have the function of reducing overlap along with signal smoothing (Agelet and Hurburgh, 2014).

Table 1. Characterization of the initial physiological quality of 10 lots of *U. decumbens* seeds.

LOT	Moisture (%)	RP (%)	NS (%)	FGC (%)	GSI	t50 (days)	SDM (g.seedling^{-1})
1	10	57 b	56 b	52 b	2.70 a	4.61 b	0.023 b
2	11	69 a	68 a	61 a	3.09 a	4.84 a	0.026 a
3	12	50 b	50 b	40 c	2.2 b	4.82 a	0.018 c
4	13	57 b	57 b	51 b	2.85 a	4.46 b	0.022 b
5	10	52 b	52 b	37 c	2.2 b	5.03 a	0.021 b
6	11	46 b	43 b	37 c	2.0 b	4.52 b	0.019 c
7	13	53 b	53 b	49 b	2.4 b	4.70 b	0.020 c
8	10	78 a	73 a	55 b	3.10 a	5.00 a	0.026 a
9	11	58 b	58 b	49 b	2.62 a	4.73 b	0.022 b
10	10	49 b	48 b	39 c	2.20 b	4.66 b	0.018 c

RP: Root protrusion; NS: Normal seedlings; FGC: First germination count; GSI: Germination speed index; t50: Time for germination of 50% of the seeds; SDM: Seedling dry matter. Means represented by the same letter do not differ from each other by Tukey test at 5% probability level.
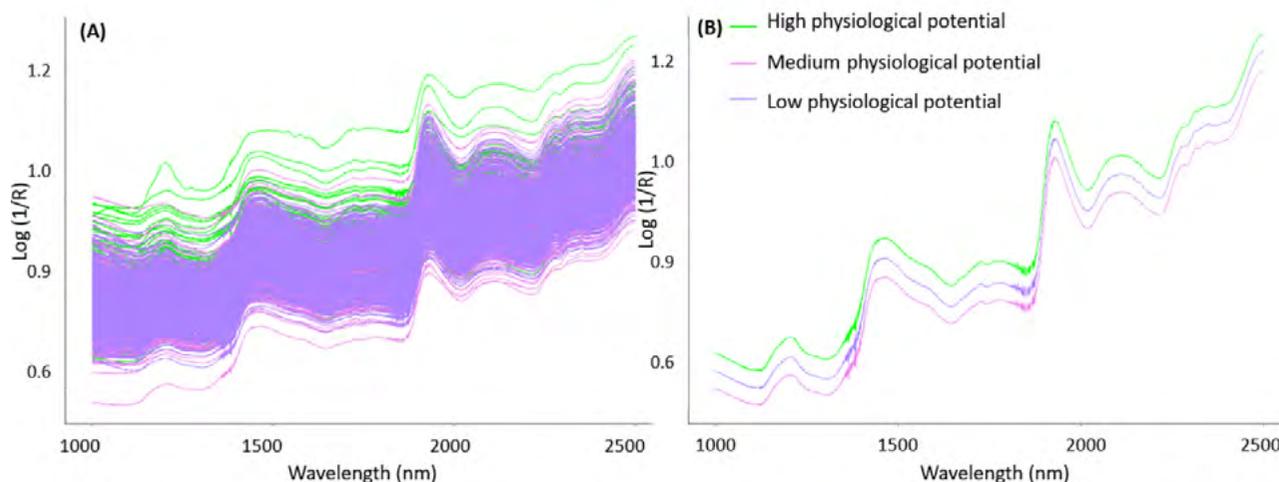

Figure 3. Original spectra (A) and mean of original spectra (B) per class of physiological potential.
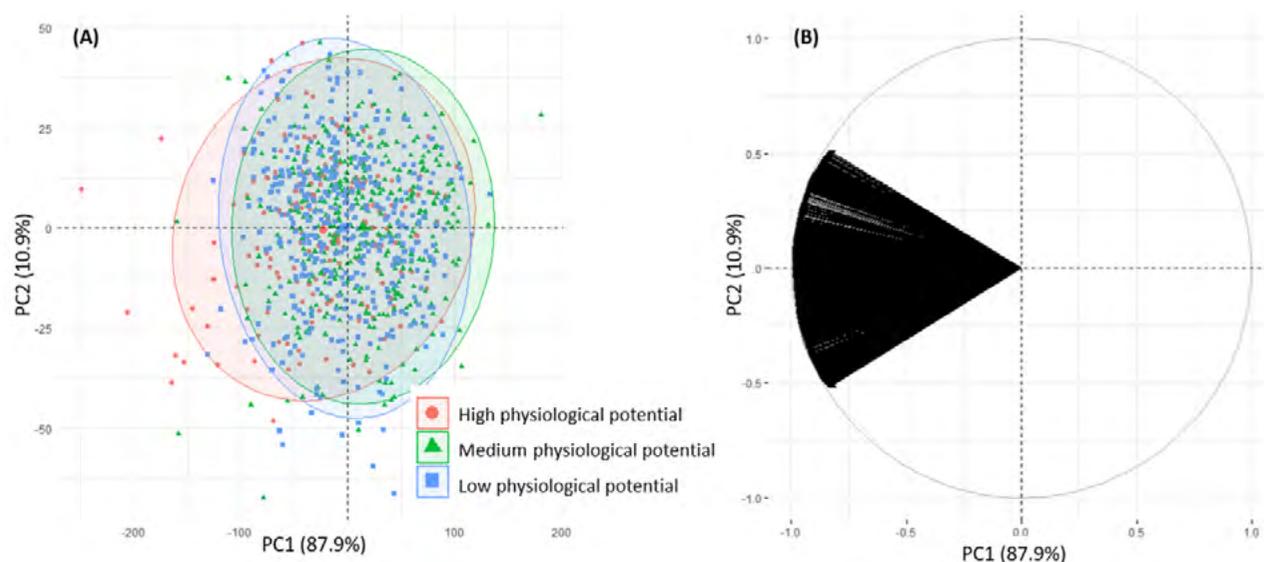
Figure 4. Principal component analysis of the original 1000 spectra with variation in physiological potential. Central ordering diagram (A) and correlation and distribution circle (B).

Table 2. Accuracy and Kappa results for training and validation (test) of the different pre-processing methods obtained by the PLS-DA classification models for *U. decumbens* seed lots. *

Pre-processing	Calibration (n*=700)		Validation (n*=300)	
	Accuracy	Kappa	Accuracy	Kappa
Original spectra	0.652	0.436	0.687	0.495
SNV[1]	0.671	0.464	0.656	0.443
MSC[2]	0.684	0.487	0.6701	0.466
1st Derivative of SG[3]	0.854	0.767	0.741	0.587
2nd Derivative of SG[3]	0.985	0.977	0.796	0.677

*n: number of spectra used for training and testing of the classification model. [1]SNV = Standard Normal Variate. [2]MSC = Multiplicative Scatter Correction. [3]SG = Savitzky-Golay.

The pre-processing recommended for the calibration model was defined based on the predictive capacity of the physiological potential classes evaluated by the results of accuracy and Kappa coefficient for the calibration and training models (Table 2). The accuracy metric is widely used to evaluate machine learning classifications, since it evaluates the proportion of the correct classification on the results obtained. However, it can be greatly influenced by the proportion of the data.

For example, in unbalanced data, it can be maximized as it considers smaller information as irrelevant (Manning et al., 2009). The Kappa coefficient is the statistical method that describes the intensity of agreement and reliability of the model, with value 1 indicating perfect agreement of the model and 0 indicating agreement equivalent to chance (Viera and Garrett, 2005).

The best classification results were obtained for the models using the Savitzky-Golay derivatives compared to the model obtained with the original data, and when applying the first derivative, the model showed accuracy of 74.1% and Kappa of 58.7% in the validation. When applying the second derivative, accuracy of 79.6% and Kappa of 67.7% were obtained in the validation, demonstrating good agreement between the results, and the pre-processing with the second derivative of Savitzky-Golay showed high efficiency to discriminate the classes in relation to the physiological potential of *U. decumbens* seeds.

The results shown for PLS-DA classification models (Table 2) with the spectral data show consistency in the data compared to other studies involving the near-infrared spectroscopy (FT-NIR) technique. This technique has also been used to classify chickpea seeds that have undergone changes due to the use of herbicides, obtaining 94% accuracy when applying the second derivative of Savitzky-Golay (Ribeiro et al., 2021). Similarly, Medeiros et al. (2022) observed that PLS-DA showed satisfactory discrimination, with a classification rate of up to 100% for NIR spectra in seeds of *Brassica* species. These authors report that NIR spectroscopy can be used to quantify the oil content and composition in seeds of these species.

Figure 5 illustrates the confusion matrix regarding the validation of the performance of the machine learning classification models for the pre-processed data with the second derivative of Savitzky-Golay, showing the predictions for the classes of high, medium and low physiological potential of the seeds. In the confusion matrix, each row of the matrix represents the instances in the actual class, whereas each column represents the instances in the predicted class. It is possible to observe that, in the calibration, there was an accuracy of almost 100% for the three classes of physiological potential (> 95.7%). In the validation, the high and medium potential classes had higher accuracy (80%), while the low potential class had 78.3%. This model is useful to prove the efficiency of applying the second derivative for treatment of the data obtained by the spectrum. In this context, it is interesting to perform an analysis in conjunction with Table 2, reinforcing the efficacy in the accuracy of the data in classifying the physiological potential of *U. decumbens* seeds.

Among the wavelengths that most contributed to the development of the PLS-DA model for the classification of *U. decumbens* seeds into different levels of physiological potential, the ranges around 1,000-1,400 nm; 1,900 nm; 2,250 and 2,400 nm, and the range around 2,500 nm stood out (Figure 6). The greatest importance for defining the model occurred at peaks in the range of 1,000 and 1,400 nm, which have harmonic bonds with the functional groups N-H, C-H and O-H, which are related to carbohydrate and protein content (Mukasa et al., 2019). Although the peaks of 1,900 nm have subtly appeared for the importance of the model, they belong to the O-H functional group associated with carbohydrates. Although the range of 2,250 to 2,400 nm was less significant, it appeared more subtly and was important for defining the model. Peaks of 2,282 and 2,330 nm are related to carbohydrates, 2,300 is related to the protein absorbance band and 2,340 is related to the C-H group of cellulose (Xu et al., 2019). Peaks in the 2,500 nm range are related to the functional groups O-H and C-H, which are also directly related to the contents of carbohydrates and CH_2 proteins (Ambrose et al., 2016; He et al., 2019).

Proteins play key roles for the growth and development of the embryonic axis, as well as for seedling formation and emergence in the field (Erbaş et al., 2016). During the seed deterioration process, there is denaturation and decrease in the content and synthesis of proteins, with protein degradation being one of the mechanisms pointed out as a cause of seed viability loss (Pinheiro et al., 2023). In this context, the deterioration process was observed in the seed lots that had lower physiological potentials, such as lots 3, 6 and 9 (Table 1).
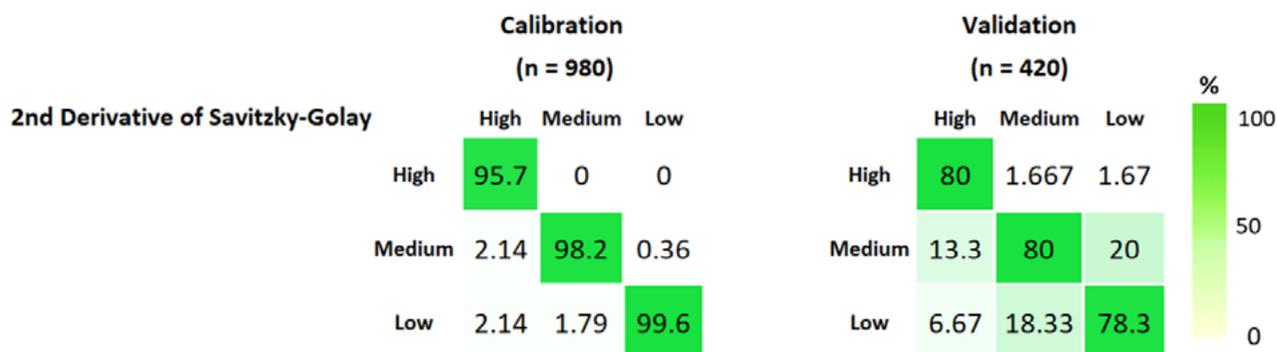

Figure 5. Confusion matrix of the classification of *U. decumbens* seed lots into vigor levels by means of the PLS-DA model using the spectra after processing with the second derivative of Savitzky-Golay.
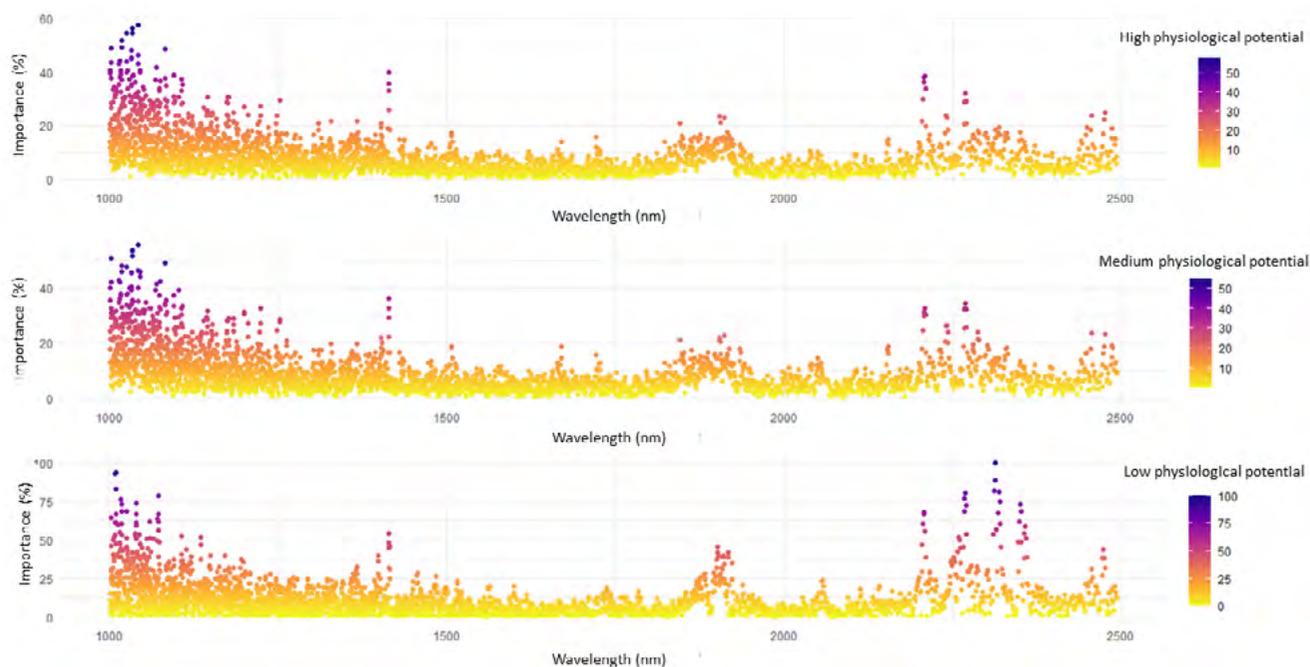
Figure 6. Importance of wavelength variables for PLS-DA classification of physiological quality levels of *U. decumbens* seeds.

As conventional biochemical analysis techniques are, in general, expensive and time-consuming, the selection of wavelengths related to chemical composition shortens the time of analysis and facilitates data processing, assisting the development of viable models for application (Orrillo et al., 2019). Thus, the results of the present study point to a series of practical applications of this methodology, which can be introduced in companies and in routine seed laboratories, research organizations and quality control programs. However, it is important to note that these methodologies do not replace the standard tests used for seed marketing, but can be an interesting tool for rapid pre-selection, and decision-making regarding seed lots.

The results obtained by ILASTIK (Figure 7) show its efficiency to classify physiological potential of *U. decumbens* seeds, being possible to identify and separate lots as to physiological potential in a similar way to what was observed by the physiological tests (Table 1). In this context, lots 3, 6 and 10 showed results indicating seeds of low physiological potential, with a higher percentage of non-germinated seeds and low percentages of strong seedlings. On the other hand, lots 2 and 8 showed higher percentages of strong seedlings, indicating the higher physiological potential of the seeds, both at 7 days (Figure 7A), and at 14 days (Figure 7B). In general, the results obtained by ILASTIK at 7 and 14 days were similar, which makes the evaluation at 7 days interesting due to the greater speed and optimization of the work (Figure 7). Therefore, it is relevant to consider that developed models can be increasingly efficient and trained, so that ILASTIK software further improves its performance, with the possibility of changing the number of images and the number of lots and treatments evaluated. ILASTIK provides all the necessary features such as probability models, graphs and features of classifiers that have a fast response and have an optimized and practical user interface for a fast interactive training experience, being able to extract accurate data through the images (Berg et al., 2019; Medeiros et al., 2020).

The PCA explained 83.1% of variability of the data obtained by ILASTIK and by the physiological tests (PC1 + PC2), reinforcing that the lots with the highest physiological potential (L2 and L8) were concentrated in the positive scores of the Component 1 (PC 1), close to the green vectors corresponding to the variables related to higher physiological quality (t50, RP, NS, GSI, SDM, FGC and strong seedlings by ILASTIK). On the other hand, the lots with

lower physiological potential (L3, L6 and L10) were concentrated in the negative scores of PC1, close to the variables that indicate lower physiological quality (weak seedlings and non-germinated seeds by ILASTIK) (Figure 8). When considering the data obtained by ILASTIK in the PCA, a high positive correlation can be observed between the vectors of green color corresponding to the variables obtained with 7 and 14 days, which confirms 7 days as a faster and more viable option for evaluation.

In summary, it was possible to correlate the data of physiological quality of *U. decumbens* seeds with those obtained by the near-infrared spectroscopy (FT-NIR) technique and seedling analysis by ILASTIK software, highlighting considerable potential for use in research, programs and at the industrial level for seed quality control and classification of lots regarding physiological potential.
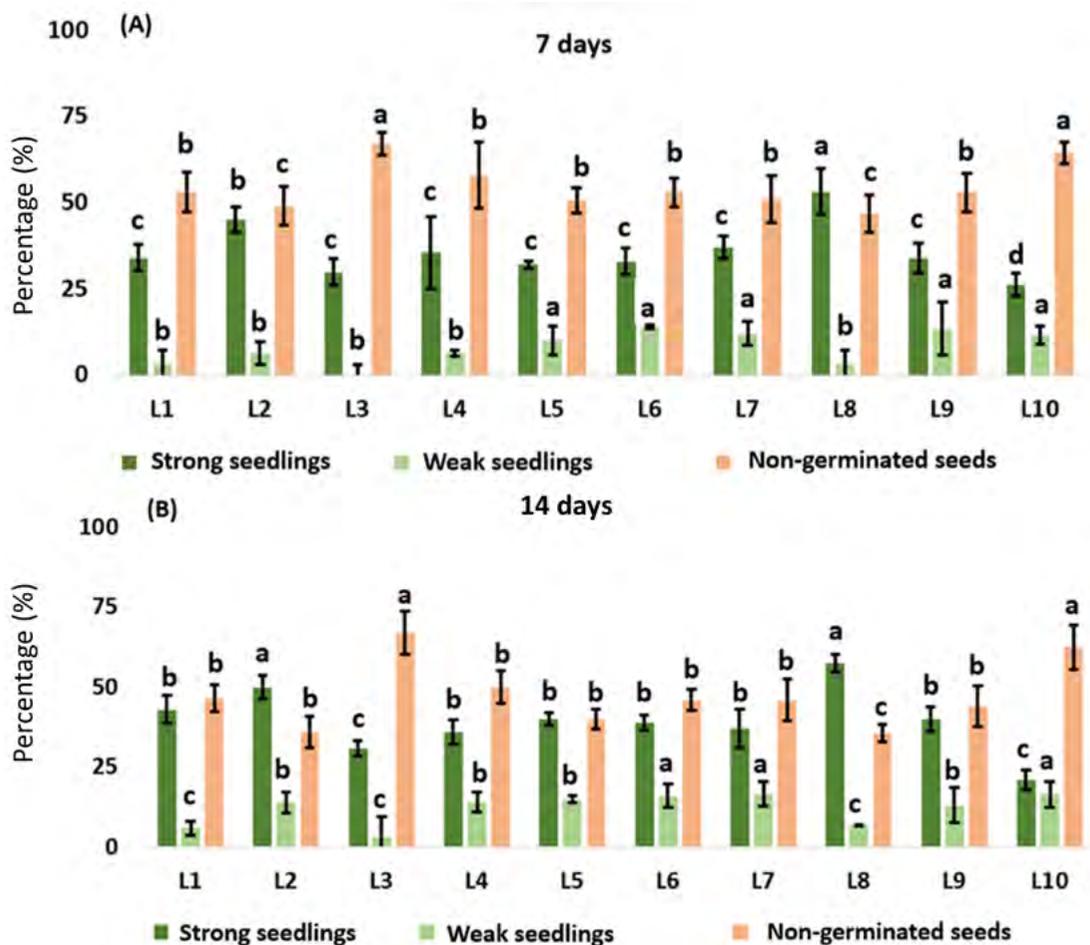

Figure 7. Percentage of strong seedlings, weak seedlings and non-germinated seeds according to the lots of *U. decumbens* seeds at 7 (A) and 14 (B) days after sowing provided by ILASTIK software. Bars represent a 95% confidence interval.
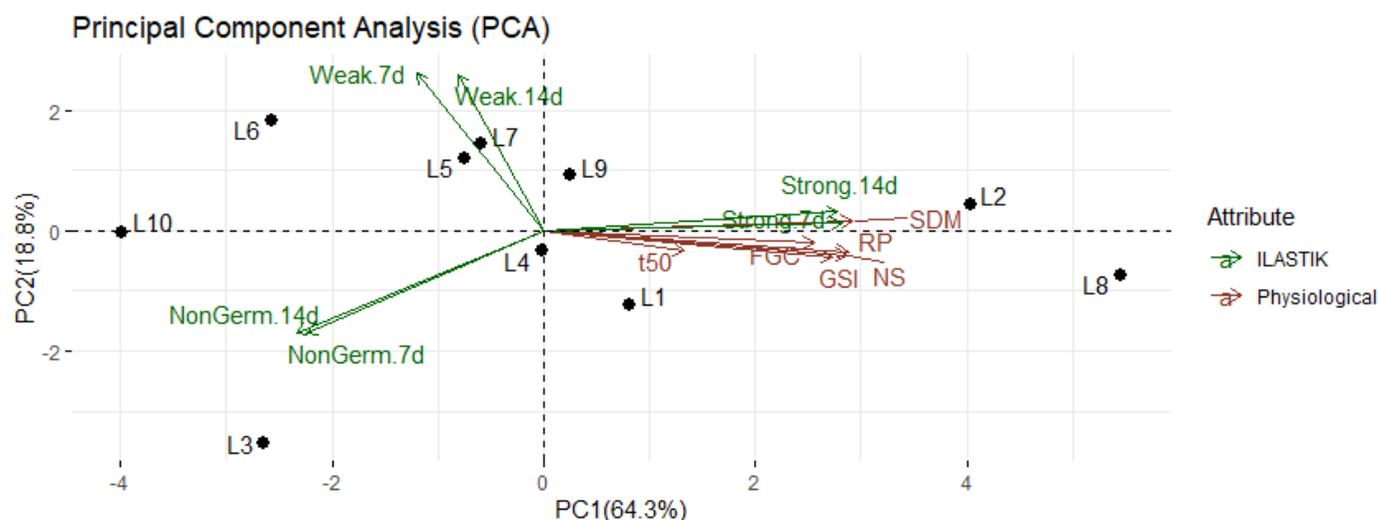
Figure 8. Principal component analysis (PCA) for the variables of characterization of the physiological quality of *U. decumbens* seeds and the data generated by ILASTIK software.

CONCLUSIONS

The FT-NIR technique combined with the PLS-DA models is sensitive to estimate, quickly and non-destructively, the physiological potential of *U. decumbens* seed lots, especially when pre-processing by means of the second derivative of the Savitzky-Golay filter is used.

Analysis of seedlings by ILASTIK software at 7 days is also efficient to evaluate the physiological potential of *U. decumbens* seeds.

ACKNOWLEDGMENTS

REFERENCES

AGELET, L.E.; HURBURGH, C.R. Limitations and current applications of Near Infrared Spectroscopy for single seed analysis. *Talanta*, v.121, p.288–299, 2014. https://doi.org/10.1016/j.talanta.2013.12.038

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. https://doi.org/10.1016/j.infrared.2015.12.008

ASMUS, G.L.; CRUZ, T. *Cultivo de Brachiaria spp. no manejo de nematoides edáficos fitoparasitos*. Empresa Brasileira de Pesquisa Agropecuária Embrapa Agropecuária Oeste. Documento 144, 2020. 22p. https://www.infoteca.cnptia.embrapa.br/infoteca/bitstream/doc/1127901/1/DOC-144.pdf

ANDRADE, G.C.; COELHO, C.M.M.; UARROTA, V.G. Modelling the vigour of maize seeds submitted to artificial accelerated ageing based on ATR-FTIR data and chemometric tools (PCA, HCA and PLS-DA). *Heliyon*, v.6, n.2, e03477, 2020. https://doi.org/10.1016/j.heliyon.2020.e03477

ANDRIAZZI, C.V.G.; ROCHA, D.K.; CUSTÓDIO, C.C. Determination of the physiological quality of corn seeds by infrared equipment. *Journal of Seed Science*, v.45, e202345002, 2023. https://doi.org/10.1590/2317-1545v45265346

BARKER, M.; RAYENS, W. Partial least squares for discrimination. *Journal of Chemometrics: A Journal of the Chemometrics Society*, v.17, n.3, p.166-173, 2003. https://doi.org/10.1002/cem.785

BATISTA, T.B.; CARDOSO, E.D.; BINOTTI, F.F.S.; SÁ, M.E.; HAGA, K.I. Nutrientes e giberelina no condicionamento fisiológico sob a qualidade de sementes de braquiária. *Revista de Agricultura Neotropical*, v.2, n.1, p.10– 16, 2015.

BERG, S.; KUTRA, D.; KROEGER, T.; STRAEHLE, C.N.; KAUSLER, B.X.; HAUBOLD, C.; SCHIEGG, M.; ALES, J.; BEIER, T.; 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, n.12, p.1226-1232, 2019. https://doi.org/10.1038/s41592-019-0582-9

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. *Regras para Análise de Sementes*. Ministério da Agricultura, Pecuária e Abastecimento. Secretaria de Defesa Agropecuária. Brasília: MAPA/ACS, 2009. 399p. https://www.gov.br/agricultura/pt-br/assuntos/insumos-agropecuarios/arquivos-publicacoes-insumos/2946_regras_analise__sementes.pdf

EMBRAPA. Empresa Brasileira de Pesquisas Agropecuárias. *Brasil cria a sua primeira cultivar do capim Brachiaria ruziziensis*. 2022. https://www.embrapa.br/busca-de-noticias/-/noticia/68876481/brasil-cria-a-sua-primeira-cultivar-de-capim-brachiaria-ruziziensis#:~:text=De%20um%20total%20180%20milh%C3%B5es,as%20braqui%C3%A1rias%20como%20g%C3%AAnero%20Urochloa.

ERBAŞ, S.; TONGUÇ, M.; ŞANLI, 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. https://doi.org/10.5073/JABFQ.2016.089.028

FERREIRA, R.C.U.; MORAES, A.C.L.; CHIARI, L.; SIMEÃO, R.M.; VIGNA, B.B.Z.; SOUZA, A.P. An Overview of the genetics and genomics of the *Urochloa* species most commonly used in pastures. *Frontiers in Plant Science*, v.12, 2021. https://doi.org/10.3389/fpls.2021.770461

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

HE, X.; FENG, X.; SUN, D.; LIU, F.; BAO, Y.; HE, Y. Rapid and nondestructive measurement of rice seed vitality of different years using near-infrared hyperspectral imaging. *Molecules*, v.24, n.12, p.2227, 2019. https://doi.org/10.3390/molecules24122227

MAGUIRE, J. D. Speed of germination - Aid in selection and evaluation for seedling emergence and vigor. Crop Science, v.2, n.2, p.176, 1962. https://acsess.onlinelibrary.wiley.com/doi/abs/10.2135/cropsci1962.0011183X000200020033x

MANNING, C.D.; RAGHAVAN, P.; SCHÜTZE, H. *Introduction to information retrieval*. Cambridge: Cambridge University Press, 2009. 544p. https://nlp.stanford.edu/IR-book/pdf/irbookonlinereading.pdf

MARCOS-FILHO, J. *Fisiologia de sementes de plantas cultivadas*. Piracicaba: FEALQ, 2015a. 660 p.

MARCOS-FILHO, J. Seed vigor testing: an overview of the past, present and future perspective. *Scientia Agricola*, v.72, n.4, p.363-374, 2015b. https://doi.org/10.1590/0103-9016-2015-0007

MARTINELLI, R.; LICERRE, R.; RUFINO, J.L.R.; CAMATARI, Y.; AZEVEDO, F.A. Roçagem ecológica com Urochloa spp.: opção de manejo integrado de plantas daninhas que promove a agricultura de conservação em citros. *Innovations Agronomiques*, v.64, p.19-29, 2018. https://web.archive.org/web/20200110205110id_/https://www6.inrae.fr/ciag/content/download/6366/46642/file/Vol64-3-Martinelli.pdf

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, 11267, 2020. https://doi.org/10.1038/s41598-020-68273-y

MEDEIROS, M.L.S.; CRUZ-TIRADO, J.P.; LIMA, A.F.; SOUZA-NETTO, J.M.; RIBEIRO, A.P.B.; BASSEGIO, D.; GODOY, H.T.; BARBIN, D.F. Assessment oil composition and species discrimination of *Brassicas* seeds based on hyperspectral imaging and portable near infrared (NIR) spectroscopy tools and chemometrics. *Journal of Food Composition and Analysis*, v.107, 104403, 2022. https://doi.org/10.1016/j.jfca.2022.104403

MUKASA, P.; WAKHOLI, C.; MO, C.; OH, M.; JOO, H.J.; SUH, H.K.; CHO, B.K. Determination of viability of Retinispora (*Hinoki cypress*) seeds using FT-NIR spectroscopy. *Infrared Physical Technology*, v.98, p.62–68, 2019. https://doi.org/10.1016/j.infrared.2019.02.008

ORRILLO, I.; CRUZ-TIRADO, J.P.; CARDENAS, A.; ORUNA, M.; CARNERO, A.; BARBIN, D.F.; SICHE, R. Hyperspectral imaging as a powerful tool for identification of papaya seeds in black pepper. *Food Control*, v.101, p.45–52, 2019. https://doi.org/10.1016/j.foodcont.2019.02.036

PAULA, L.C.; CEZÁRIO, A.S.; OLIVEIRA, N.C.; LIMA, M.; VIEIRA, J.P.B.; DAMASCENA, E.G.; RABELO, A.P.D.; SANTOS, W.B.R. Manejo e adaptação de pastagens do gênero *Urochloa* em solos do cerrado. *Colloquium Agrariae.* v.13, p.276-288, 2017. https://doi.org/10.5747/ca.2017.v13.nesp.000233

PINHEIRO, D.T.; DIAS, D.C.F.S.; SILVA, L.J.; MARTINS, M.S.; FINGER, F.L. Oxidative stress, protein metabolism, and physiological potential of soybean seeds under weathering deterioration in the pre-harvest phase. *Acta Scientiarum. Agronomy*, v.45, e56910, 2023. https://doi.org/10.4025/actasciagron.v45i1.56910

R CORE TEAM. *R. A Language and Environment for Statistical Computing*, 2022.

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.S. FT-NIR and linear discriminant analysis to classify chickpea seeds produced with harvest aid chemicals. *Food Chemistry*, v.342, e.128324, 2021. https://doi.org/10.1016/j.foodchem.2020.128324

RIBEIRO, M.R.; SIMEONE, M.L.F.; TRINDADE, R.S.; DIAS, L.A.S.; GUIMARÃES, L.J.M.; TIBOLA, C.S.; AZEVEDO, T.C. Near infrared spectroscopy (NIR) and chemometrics methods to identification of haploids in maize. *Microchemical Journal*, v.190, 108604, 2023. https://doi.org/10.1016/j.microc.2023.108604

RINNAN, A.; BERG, F.V.D.; ENGELSEN, S.B. Review of the most common pre-processing techniques for near-infrared spectra. *Trends in Analytical Chemistry,* v.28, n.10, p.1201–1222, 2009. https://doi.org/10.1016/j.trac.2009.07.007

SILVA, L.J.D.; MEDEIROS, A.D.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. https://doi.org/10.1590/2317-1545v42n2217267

VIERA, A.J.; GARRETT, J.M. Understanding interobserver agreement: the kappa statistic. *Family Medicine*, v.37, n.5, p.360–363, 2005. http://www1.cs.columbia.edu/~julia/courses/CS6998/Interrater_agreement.Kappa_statistic.pdf

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. https://doi.org/10.1016/j.aiia.2019.05.001

XU, J.; NWAFOR, C.C.; SHAH, N.; ZHOU, Y.; ZHANG, C. Identification of genetic variation in *Brassica napus* seeds for tocopherol content and composition using near-infrared spectroscopy technique. *Plant Breeding*, v.138, n.5, p.624–634, 2019. https://doi.org/10.1111/pbr.12708