

Use of near infrared spectroscopy in cotton seeds physiological quality evaluation

Lívia Giro Mayrinck^{1*} , Juliana Maria Espíndola Lima¹ , Gabriel Castanheira Guimarães¹ , Cleiton Antônio Nunes¹ , João Almir Oliveira¹ 

ABSTRACT: This study aimed to evaluate the near-infrared spectroscopy potential in analyzing the quality of cottonseed regarding different physiological quality levels, noting the need for faster techniques and tools to aid decision making. It was used eight samples of cottonseed with and without lint, presenting different physiological quality. The “high” (lots 1, 4, 5, 6 and 7) and “low” (lots 2, 3 and 8) vigor levels were defined based on vigor tests carried out and on the Normative Instruction 45/2013. The near infrared spectroscopy spectra was obtained from four types of sample preparations: whole seeds, cut in a half, without tegument and grounded seeds. Using the spectra and the grouping of lots in high and low vigor, cross validation models were optimized, built using the PLS - DA method, making it possible to predict seed classes. Grounded seeds were the best type of sample preparation, with 95% of correct predictions for high vigor seeds and 100% of low vigor (both for seeds with lint) and with 100% correct predictions for high vigor seeds and 91.7% low vigor (without lint).

Index terms: *Gossypium hirsutum* L., vigor, chemometrics.

Uso da espectroscopia no infravermelho próximo na avaliação da qualidade fisiológica de sementes de algodão

RESUMO: Este trabalho foi conduzido com o objetivo de avaliar o potencial da espectrometria no infravermelho próximo para analisar a qualidade de sementes de algodão em função de diferentes níveis de vigor, visto a necessidade por técnicas e ferramentas mais rápidas para o auxílio na tomada de decisão. Foram utilizadas oito amostras referentes a oito lotes de semente de algodão, com e sem línter, de qualidades fisiológicas diferentes. Foram definidos os níveis “alto” (lotes 1, 4, 5, 6 e 7) e “baixo” (lotes 2, 3 e 8) de vigor, baseados em diferentes testes de vigor realizados e na Instrução Normativa 45/2013. Os espectros obtidos pelo equipamento de espectroscopia no infravermelho próximo foram provenientes de quatro preparos de amostra: sementes inteiras, cortadas, sem tegumento e maceradas. Utilizando-se dos espectros gerados e o agrupamento dos lotes em níveis, foram otimizados modelos de validação cruzada, construídos a partir do método PLS – DA, predizendo as classes de sementes. O melhor preparo de amostra foi o de sementes maceradas, com 95% de predições corretas para sementes de alto vigor e 100% das de baixo vigor (com línter) e com 100% de predições corretas para sementes de alto vigor e 91,7% de baixo vigor (sem línter).

Termos para indexação: *Gossypium hirsutum* L., vigor, quimiometria.

Journal of Seed Science, v.42,
e202042016, 2020

<http://dx.doi.org/10.1590/2317-1545v42227169>

*Corresponding author
E-mail: lgmayrinck@hotmail.com

Received: 10/21/2019.
Accepted: 4/15/2020.

¹Agrônomos, Universidade Federal de Lavras, Campus Universitário, Caixa Postal 3037, 37200-000, Lavras, Minas Gerais, Brasil,

INTRODUCTION

The cotton crop represents an important part of generating income and employment on a national and global scale. Its cultivation is mainly destined to the production of fiber, in which, the harvest of 2019/2020 the production estimate was about 2,853.7 thousand tons of cotton plume. The total area planted in Brazil was 1,670.80 thousand hectares, which is the largest of the last five harvest years (CONAB, 2020).

Considering all the technology used for cotton production, it is important to use high quality seeds in order to obtain satisfactory results in the field. Mattioni et al. (2012) observed that the use of seeds of medium and low quality originate plants that cannot be equated, in development and productivity, with those of high-quality seeds.

In the post-harvest stages of cottonseed, the evaluation of the seed lots allows estimating their aggregate value for both reception, delineation and commercialization purposes, as well as to predict the number of seeds required for planting. Thus, vigor tests that differentiate the physiological potential of the materials are carried out so that complementary information is obtained for the internal quality control of the companies producing cottonseed, but are generally, they are more time-consuming evaluations.

Seeking faster responses with consistent results with the actual quality of cottonseed, the near-infrared spectroscopy combined with chemometric methods can represent a promising alternative for such information.

The advantages of this equipment are the ability of the analyzes to be carried out successively, in a short period of time. Capable of generating a large quantity of information, with less need of labor, speed, less cost, not polluting, does not use chemicals or reagents and it can be non-destructive (Amorim, 1996).

Seed research has been successful in evaluating quality, such as soybeans, cotton and tomato (Bazoni et al., 2017; Gaitán-Jurado et al., 2008; Huang et al., 2013; Shrestha et al., 2016), viability in maize and spinach (Ambrose et al., 2016; Shetty et al., 2012), composition and prediction of phytic acid in *Vigna radiata* (Pande and Mishra, 2015), nitrogen content in *Vigna unguiculata* (Ishikawa et al., 2017), oil composition in sunflower and canola (Grunvald et al., 2014; Rossato et al., 2013), as well as the classification of genotypes in cotton, barley, castor beans and maize (Cui et al., 2012; Jia et al., 2015; Ringsted et al., 2016; Santos et al., 2014; Soares et al., 2016).

This research aimed to evaluate the potential of the near-infrared spectroscopy using different seed sample procedure to determine the physiological quality of cottonseed with and without lint.

MATERIAL AND METHODS

The experiment was conducted at the Central Laboratory of Seed Analysis, and in the Laboratory of Seed Pathology of the *Universidade Federal de Lavras – UFLA*, Lavras, Minas Gerais. The same was conducted in two parts. In the first, physiological tests were carried out to determine the physiological quality levels of cottonseed samples. In the second, the samples were analyzed in the near-infrared spectroscopy equipment for the detection of these different levels of cotton seed quality.

The seeds used were supplied by the company Bayer®, in which these samples were from eight lots of different physiological seed qualities, in which each sample was divided into seeds with and without lint, and to identify each sample it was given a numbering from 1 to 8.

Physiological tests

The experimental design used to determine the physiological quality of the cottonseed was completely randomized with factorial 2 x 8 (with and without lint x samples) and four replications.

At the reception of the samples of cottonseed were evaluated according to:

Germination: four replications of 50 seeds, sow on a roll of Germitest® paper moistened 2.5 times the weight of the substrate in distilled water and kept in a germinator at 25 °C, a single count was carried out at 5th day after sowing, counting the normal seedlings higher than 5 cm. The results were expressed as a percentage (Brasil, 2009a).

Seedling emergence: four replications of 50 seeds per treatment were used under uncontrolled environmental conditions. The substrate of the bed was composed of sand and soil in a ratio of 2:1. The counting of emerged seedlings was made on the 12th after sowing. The results were expressed as a percentage.

Seedling stand: conducted together with the seedling emergence test, counting was registered on the 7th and 21st day after sowing, with the results expressed as a percentage of the number of emerged seedlings on the respective counting days.

Speed of seedling emergence: also conducted with the seedling emergence test, where the number of seedlings emerged was counted daily until day 14th, using the counting parameter of Maguire (1962).

Tetrazolium: four replications of 50 seeds were used for each treatment. The seeds were soaked between germitest paper moistened with distilled water at 25 °C for 16 hours. Then the seed tegument was removed and stained in 0.075% tetrazolium salt for 4 hours at 30 °C, in the absence of light. The results were expressed as a percentage of vigorous and viable seeds, according to Vieira and Von Pinho (1999).

Seed health: eight replications of 25 seeds arranged on two sheets of filter paper were used in Petri dishes, then 2.4-D, distilled water, and agar were added for soaking the filter paper. The Petri dishes were incubated seven days in a chamber for 12 hours at 20 °C. Each seed was analyzed individually with the aid of a microscope to identify the incidence of fungus. The results were expressed in percentage. The filter paper method was made, as described in Brasil (2009b).

The data were submitted to analysis of variance by the SISVAR software and the means were compared by the Tukey test ($p \leq 0.05$) (Ferreira, 2011), except for the sanity test, due to the large data variation.

Based on the results of the physiological tests, the quality of the cotton samples was classified using the marketing standard used by the *Ministério da Agricultura Pecuária e Abastecimento* (MAPA), dated 09/17/2013 -IN 45/2013 (Brasil, 2013). Where cottonseed to be commercialized, must be without lint and have at least 75% of germination. Thus, two levels of physiological quality (high and low) were considered, being then, samples with values above 75% of germination considered of high quality and below this value of low quality. This classification was performed for later comparison with the data obtained in the analysis of near-infrared spectroscopy.

Near-Infrared Spectroscopy

For each sample of cottonseed with and without lint, four types of seed preparation were tested:

Whole seed: did not have any pre-treatment, which was exposed directly to the output of the infrared light. The number of seeds used was 100 units, each representing one replication.

Seed cut in half: each seed was sectioned longitudinally with a scalpel and exposed to the infrared light, where the seed cut band was faced the light right after cutting to avoid oxidation. The number of seeds used was 100 units, each representing one replication.

Seed without tegument: the tegument was removed from the seed using a scalpel and exposed directly to the infrared light right after the procedure to avoid oxidation of the sample. The number of seeds used was 100 units, each representing one replication.

Grounded seed: a mill was used, with the addition of liquid nitrogen and polyvinylpyrrolidone into 40 seeds per replication, to avoid oxidation of the sample, and then the powder was placed in cuvettes for exposure to the infrared light. It was used 20 replications in the total of each treatment.

Near-infrared Spectroscopy: the spectra were obtained by placing the samples directly or buckets (ground seeds), at the output of the infrared source of the tensioner device 27 Bruker®, generating spectra by FT-IR detector (Fourier-Transform Near-Infrared) coupled with the help of OPUS_ Spectroscopy software version 6 from the same equipment manufacturer. To compile the reading database, the spectrometer collected 48 scans at each measurement of absorbance, with a resolution of 8 cm⁻¹, in the range 10000 to 4000 cm⁻¹, per replication.

Multivariate analysis: a cross-validation model was optimized from the grouping of the cottonseed samples in high and low quality, using 3/4 of the samples for calibration and 1/4 for the test, constructed from the method of multivariate classification by partial least squares regression with discriminant analysis (PLS-DA) with multiplicative

scatter correction preprocessing of the data. The wavelength between 8000 cm⁻¹ and 10000 cm⁻¹ were removed from the analysis due to noise. It was used the Pirouette® statistical software, with the Y classes being the dependent variables and the obtained spectra constituting the independent variables X (Abdi, 2003).

The sensitivity and specificity of the optimized models were obtained by dividing the number of predicted samples of the class by the total number of predictions of the class, and dividing the number of samples predicted as not being of the class by the total number of samples that are not of the class, respectively (Szymanska et al., 2011), in the validation stage.

RESULTS AND DISCUSSION

The variance analysis of the physiological tests showed significant differences between the cottonseed samples studied, which was possible to determine the two levels of seed quality desired for this work, as high and low physiological quality levels.

The results of germination (Table 1) for cottonseed with lint presented better quality to the samples one, two, four, and six; otherwise, the samples three, five, seven, and eight showed lower quality. Cottonseed without lint had a higher percentage of germination for samples one, four, five, six, and seven compared to samples two, three, and eight. The presence and absence of the lint in the seed could affect the physiological quality in different ways. Seeds with lint can have low quality when the lint makes the perfect media for fungus proliferation (samples four, five and seven); otherwise, it can also help to protect the seed from fungus as a barrier (samples one, two and six).

Seeds without lint can have a decrease of germination if the sulfuric acid treatment passes through the tegument damaging the embryo, and this could happen if the seed has physical damage or the sulfuric acid damages the tegument (samples two and three). Thus, if the treatment is successful in only removing the lint, percentages of germination can be increased (samples four, five, six and seven), besides making it easier for seed sow in the field. According to the results of cottonseed germination between the treatments with and without lint, the authors Queiroga et al. (2009) found that seeds without lint had better physiological qualities.

For the commercialization of cottonseed, it is required that the seeds are without lint and present above 75% germination, as indicated in the Normative Instruction 45 of Brazil's Ministry of Agriculture – (Brasil, 2013), which is routinely used by the cottonseed producing companies. From this premise, it was defined based on the germination

Table 1. Means of germination (GERM, %) of samples 1 to 8 of cotton seeds, with and without lint.

Samples	GERM	
	With	Without
1	95 aA*	94 aA
2	86 abA	74 bB
3	56 deA	46 cB
4	81 bcB	93 aA
5	74 cB	91 aA
6	90 abA	95 aA
7	61 dA	86 aA
8	45 eA	46 cA
C.V. (%)**	6.94	
Mean (%)	75.93	

*Means followed by the same letter, upper case in the line and lower case in the column, do not differ by Tukey test, 5% of probability.

**C. V. – coefficient of variation.

test (Table 1) of cottonseed without lint that samples one, four, five, six and seven were eligible for commercialization, and samples two, three and eight did not fit in this category, which was used as a parameter to determine high and low-quality seed samples.

In the evaluation of seedling emergence (Table 2) from seed with lint, it was possible to observe superiority for sample 1 and inferiority for samples four, seven and eight, with the others presenting intermediate values of vigor. In seeds without lint samples, three and eight had low vigor compared to the others. When it was compared the vigor of seeds with and without lint, only sample two lowered vigor after the delinting.

The speed of seedling emergence (Table 2) seedlings with lint showed to be similar to most of the samples, differentiating the sample one with the highest emergence speed and the sample eight the lowest. For seeds without lint, there was more significant variation between samples, but sample one remained with higher value, samples three and eight as inferior, and the others presented intermediate values for this variable.

The results of the stand (7 and 21 days) for cotton seeds without lint were similar to those observed for germination, with samples three and eight having the lower stand, which differed statistically from the others; the stand (7 and 21 days) of seeds with lint had sample eight showing the lower percentage.

Observing the results obtained from seedling emergence in the stand of 7 and 21 days (Table 2), it could not be seen any relevant variation from the seventh to the twenty-first counting day. It could be inferred that the physiological quality and the presence or absence of lint did not influence the establishment of the seedling stand.

Comparing the results of germination, seedling emergence, speed of seedling emergence and stand of 7 and 21 days, it could be seen that sample eight presented inferiority in all the tests, for presence or absence of lint. These tests were also able to differentiate the sample three as inferior physiological quality, except when evaluated seed with lint (Table 2).

However, sample two showed no differentiation from the samples with higher vigor. That might have occurred because this sample has intermediate vigor between high and low-quality seeds, with no differences detected in these tests (Steiner et al., 2011).

The last evaluation made it identify the quality of the eight cottonseed samples was the tetrazolium test, and cotton seeds with lint presented for most of it higher or similar percentages of vigorous and viable seeds compared to the seeds without lint (Table 3).

In the seed production system, cotton seeds undergo a long period of processing. The physical processes that these

Table 2. Mean results of seedling emergence (SE, %), speed of seedling emergence (SSE, index), stand of 7 and 21 days (%) of cottonseed obtained from different samples, with and without lint.

Samples	SE		SSE		stand 7		stand 21	
	With	Without	With	Without	With	Without	With	Without
1	91.0 aA	94,5 aA	10.22 aB	14.30 aA	88 aA	94 aA	90 aA	94 aA
2	82.0 abA	80.0 aA	8.98 abB	11.77 abA	80 aA	79 aA	81 abA	80 aA
3	84.0 abA	55.5 bB	9.35 abA	7.65 cdA	82 aA	53 bB	84 abA	55 bB
4	74.0 bcB	88.5 aA	8.04 abB	11.21 bA	73 abB	87 aA	74 bcB	90 aA
5	75.0 abcB	95.0 aA	7.88 abB	12.68 abA	74 abB	92 aA	74 abcB	94 aA
6	83.5 abA	83.0 aA	8.98 abA	10.03 bcA	81 aA	81 aA	83 abA	85 aA
7	72.0 bcB	83.3 aA	8.52 abB	11.81 abA	72 abA	81 aA	72 bcA	81 aA
8	62.0 cA	56.0 bA	6.35 bA	6.43 dA	60 bA	50 bA	65 cA	54 bB
C.V. (%)	9.26		14.07		9.88		8.99	
Mean (%)	78.7		9.64		76.89		78.73	

*Means followed by the same letter, upper case line and lower case in column, do not differ by Tukey test, 5% probability.

Table 3. Mean results of tetrazolium evaluated according to vigorous and viable (%) cottonseed, obtained from different samples, with and without lint.

Samples	Vigorous		Viable	
	With	Without	With	Without
1	88 abA	87 aA	97 aA	95 abA
2	81 abcA	59 bB	95 aA	90 abA
3	88 abA	52 bB	96 aA	79 cB
4	62 dB	80 aA	91 aA	95 aA
5	79 bcA	83 aA	95 aA	98 aA
6	94 aA	85 aA	99 aA	97 aA
7	92 abA	89 aA	97 aA	98 aA
8	73 cdA	58 bB	90 aA	84 bcA
C.V. (%)	8.18		5.5	
Mean (%)	77.93		93.31	

*Means followed by the same letter, upper case in the line and lower case in the column, do not differ by Tukey test, 5% probability.

seeds undergo (harvest, fiber withdrawal, delinting) cause irreversible damages, reducing the physiological quality that can be detected in germination and vigor tests. The tetrazolium can be used in different stages of the processing, which make it able to monitor the origin of the damage, the severity and its extension, representing an important tool for the evaluation of the cottonseed quality (Mattioni et al., 2012; Zorzal et al., 2015).

As for the separation of the vigorous seed samples without lint, the same results were observed as the germination test and similarities with the other tests, samples two, three and eight were statistically inferior to the others. The mechanical damages represented the main causes of the low vigor of the evaluated seed with and without lint. In the results of viable seeds, similarities to the other tests were also observed, as samples three and eight having the lower means.

Due to the high amplitude of the seed health test, the data was not presented, no correlation that could have influence directly physiological quality was found.

The fungus incidence in the seeds with and without lint was: *Colletotrichum gossypii*, *Fusarium oxysporum*, *Penicillium* sp., and *Aspergillus* sp. The presence of pathogens in the seeds implies a lower health quality of the seeds and may impair the storage and performance of the plants in the field (Gama et al., 2012; Pedroso et al., 2010).

According to the results of the physiological tests, basing the NI 45 of cottonseeds commercialization, the samples were classified in high (samples one, four, five, six and seven) and low quality (samples two, three and eight) to be used as a parameter for the near-infrared spectroscopy test.

Means for NIR spectra from samples of high and low physiological quality, with and without lint, are shown in Figures 1 and 2. The data were compared by using multiplicative scatter correction (MSC), using the Chemoface® software (Nunes et al., 2012).

Molecular absorptions in NIR spectroscopy are not intense, which may result in overlapped bands, making them wide, lowering the sensitivity, and making it necessary to use calibration models. Thereafter, statistical tools may be associated with chemical data. The PLS-DA method used seeks the linear relationship between the dependent variables "Y" (estimation of the variables of interest) and independent variables "X" (spectra), through the regression by partial least squares (PLS), which creates correlations between similarities and structural differences between the compounds, allowing the interpretation of a series of complex data with a large number of variables.

Although only minor visual differences for spectra means were observed (Figures 1 and 2), both for high and low physiological quality, such difference were detected by the models. Furthermore, spectra means were similar for cotton seed samples with and without lint from the same preparation type, as well as among different preparation types

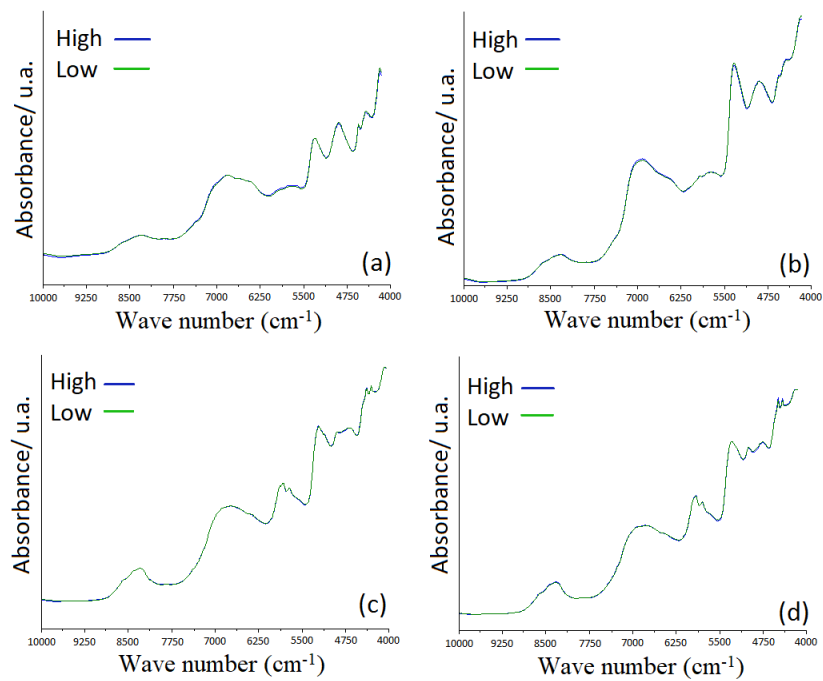


Figure 1. Spectra in near-infrared region corresponding to different preparation types of samples after mean and application of multiplicative scatter correction preprocessing of samples with and without lint to high and low physiological quality. (a) Whole cottonseeds with lint, (b) Whole cottonseeds without lint, (c) Cut in a half cottonseed with lint, (d) Cut in a half cottonseed without lint. The x-axis corresponds to the wave number (cm^{-1}) and the y-axis to absorbance (u.a.^{-1}).

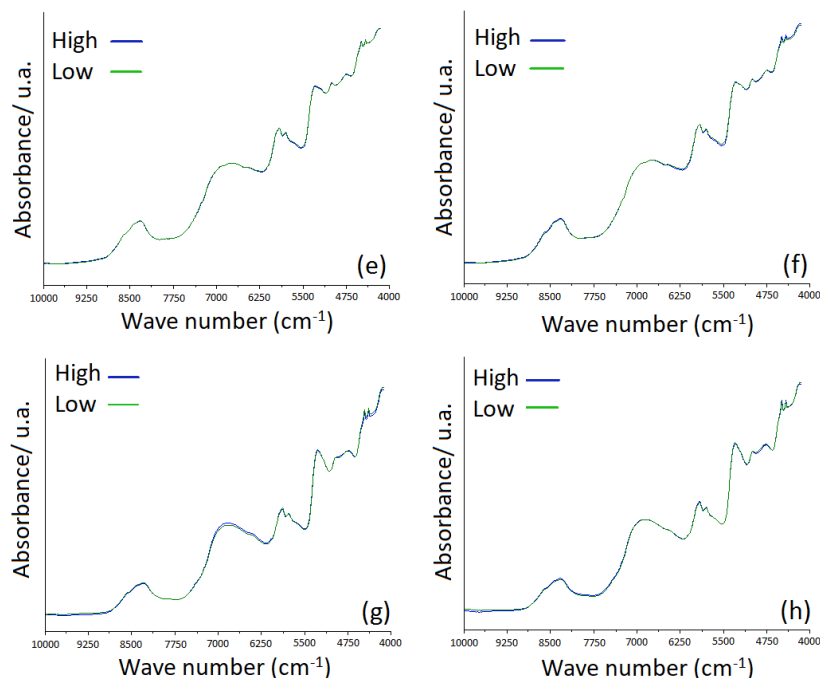


Figure 2. Spectra in near-infrared region corresponding to different preparation types of samples after mean and application of multiplicative scatter correction preprocessing of samples with and without lint to high and low physiological quality. (e) Cottonseeds without tegument with lint, (f) Cottonseeds without tegument without lint, (g) Grounded cottonseeds with lint, (h) Grounded cottonseeds without lint. The x-axis corresponds to the wave number (cm^{-1}) and the y-axis to absorbance (u.a.^{-1}).

(whole seeds, cut in a half, without tegument and grounded seeds).

As a supervised pattern recognition method, PLS-DA requires some type of information to build the model in advance, thus the “high quality” and “low quality” classes were added as the dependent variables (y) for the construction of the calibration model.

In Table 4 is presented multivariate analyses of the near-infrared spectroscopy test, and better results were obtained for the grounded seeds with and without lint, which separated samples with a high percentage. This type of sampling method has a larger contact surface to the penetration of the radiation light to the compounds of the integument and embryo at the same time, representing a greater quantity of sources of discrimination, for better predictions between the different seed quality. This type of sampling method also homogenizes the analytical material, since even in good seed sample; there may be lower quality seeds, which could influence the result.

The models obtained for the grounded seeds with lint (Table 4), reached 95% and 100% of correct answers in the external validations for high and low quality, respectively. It is important to observe the results of the external validation, since it indicates the predictive capacity of the model for seeds that did not participate in the construction of the same, indicating that it is robust and capable of evaluating different samples.

Models for the high-quality grounded seeds without lint it was obtained 92.3% of hits answers, 4.6% of error and 3.1% of seeds not recognized in the cross-validation, otherwise, for the external validation, there was 100% of correctness. For the low-quality samples, there was 97.4% accuracy in the cross-validation and 91.7% in the external validation.

In the evaluation of the performance of the optimized models, values of sensibility and specificity had 1.00 in both values for a model (Table 4); for Forina et al. (1991), this is a perfect model.

In the sampling method that used whole seeds for the modeling, it was observed a great percentage of seeds

Table 4. Mean percentages of hits and error in the crosses and external validations, and values of sensibility and specificity of the optimized models for seed sampling (whole, cut in half, without tegument and ground) of cotton seeds, with and without lint, correlated with the high (samples 1, 4, 5, 6 and 7) and low (samples 2, 3, and 8) physiological quality classification using the technique of near-infrared spectroscopy, with the aid of PLS-DA.

Seed Sampling	Sample Quality	Cross-validation			External validation			Sensibility	Specificity
		Hits		Error	Hits		Error		
		High	Low		High	Low			
Whole (with)	High	86.9	10.7	2.4	69.6	25.6	4.8	0.73	0.71
	Low	26.7	71.1	2.2	25.3	74.7	0.0		
Whole (without)	High	95.7	3.5	0.8	92.0	8.0	0.0	0.83	0.91
	Low	5.3	94.2	0.4	15.3	82.4	2.4		
Cut in half (with)	High	86.7	11.2	2.1	76.0	20.0	4.0	0.56	0.64
	Low	20.0	76.4	3.6	54.7	42.7	2.7		
Cut in half (without)	High	86.1	10.7	3.2	85.6	10.4	4.0	0.71	0.82
	Low	20.9	75.6	3.6	27.6	65.8	6.6		
Without tegument (with)	High	80.8	12.0	7.2	49.6	38.4	12.0	0.58	0.56
	Low	30.2	60.0	9.8	21.3	65.3	13.3		
Without tegument (without)	High	92.3	6.4	1.3	84.8	12.8	2.4	0.63	0.76
	Low	13.3	80.4	6.2	44.0	49.3	6.7		
Ground (with)	High	89.2	7.7	3.1	95.0	5.0	0.0	1.00	0.95
	Low	10.3	84.6	5.1	0.0	100.0	0.0		
Ground (without)	High	92.3	4.6	3.1	100.0	0.0	0.0	0.92	1.00
	Low	2.6	97.4	0	8.3	91.7	0		

classified correctly. Similar the sampling method of grounded seeds. The separation of seed samples with lint obtained lower percentages of accuracy compared to seed the samples without lint, both for high and low-quality samples (Table 4). This difference may have occurred when the lint influenced the scattering of the infrared lights into the sample, acting as a thin physical barrier for the penetration of the radiation.

Using the near-infrared spectroscopy allied to chemometrics, Soares et al. (2016) classified cotton seed four high genetic quality cultivars (no lint) by PLS-DA method, obtaining 96.91% correctly classified cultivars at the first validation, and 88.66% in the second validation. Thus, it is observed a potential use of the whole cottonseed without lint for the classification between different parameters by the PLS-DA method allied to NIR spectroscopy.

The values of sensibility and specificity obtained in the models optimized for the whole cottonseed were more dissimilar than the grounded seed values when comparing the results of the seeds with and without lint.

From the cut in half cottonseed, less efficient models were obtained than those previously presented. It was observed the occurrence of a super fit model for the cut in half seed of low vigor with lint, as it presented good results for the samples belonging to the calibration set, but failed to predict external samples (Table 4). The values of sensitivity and specificity obtained in these optimized models were also more discrepant comparing the results of the seeds with and without lint. Oxidation of the seed after the cut could have been the main issue in this sampling method.

The models generated for cottonseed without tegument did not present a high percentage of correct answers. There was disagreement between the classifications of the high and low-quality samples, counting on super adjusted models and the lower values of sensitivity and specificity obtained among the different modes of preparation. The damages caused when the tegument was removed could be an issue making such a large variation for the validations.

CONCLUSIONS

The chemometric analysis (PLS-DA), allied to near-infrared spectroscopy, made it possible to predict the physiological quality of cottonseed with good accuracy, representing a promising technique.

The models optimized for grounded cottonseed, with and without lint, promote the percentage of correctness and higher values of performance among the other models.

REFERENCES

- ABDI, H. Partial least square regression (PLS regression). *Encyclopedia for research methods for the social sciences*, v. 6, n. 4, p. 792-795, 2003.
- AMBROSE, A.; LOHUMI, S.; LEE, W. H.; CHO, B. K. Comparative nondestructive measurement of corn seed viability using Fourier transform near-infrared (FT-NIR) and Raman spectroscopy. *Sensors and Actuators B: Chemical*, v. 224, p. 500-506, 2016. <https://www.sciencedirect.com/science/article/abs/pii/S0925400515305475>
- AMORIM, H. V. *Manual de métodos analíticos para o controle da produção de álcool e açúcar*. 2. ed. Piracicaba: Esalq-USP, 1996. 194p.
- BAZONI, C. H.; IDA, E. I.; BARBIN, D. F.; KUROZAWA, L. E. Near-infrared spectroscopy as a rapid method for evaluation physicochemical changes of stored soybeans. *Journal of Stored Products Research*, v. 73, p. 1-6, 2017. <https://www.sciencedirect.com/science/article/abs/pii/S0022474X17301315>
- BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. *Instrução Normativa nº 45*, de 17 de setembro de 2013. Diário Oficial da República Federativa do Brasil, Brasília, DF, n. 183, 20 set. 2013. Seção 1, p. 6-27. http://www.lex.com.br/legis_24861657_INSTRUCAO_NORMATIVA_N_45_DE_17_DE_SETEMBRO_DE_2013.aspx
- 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. 395p. http://www.agricultura.gov.br/arq_editor/file/2946_regras_analise_sementes.pdf

- BRASIL, Ministério da Agricultura, Pecuária e Abastecimento. Testes de sanidade: Incubação em substrato de papel ou método do papel de filtro "blotter test". *Manual de análise sanitária de sementes*. Brasília: MAPA/ACS. ISBN 978-85-99851-64-7. Pt 5, p. 33-34, 2009b. 200 p. <https://www.abrates.org.br/files/manual-de-analise-sanitaria-de-sementes.pdf>
- CONAB. Companhia Nacional de Abastecimento. <https://www.conab.gov.br/info-agro/safras/serie-historica-das-safras>. Accessed in 17th Mars 2020.
- CUI, H. F.; YE, Z. H.; XU, L.; FU, X. S.; FAN, C. W.; YU, X. P.. Automatic and rapid discrimination of cotton genotypes by near infrared spectroscopy and chemometrics. *Journal of Analytical Methods in Chemistry*, v. 2012. <https://www.hindawi.com/journals/jamc/2012/793468/>
- FERREIRA, D. F. Sisvar: a computer statistical analysis system. *Ciência e Agrotecnologia*, v. 35, n. 6, p. 1039-1042, 2011. http://www.scielo.br/scielo.php?pid=S1413-70542011000600001&script=sci_arttext
- FORINA, M.; ARMANINO, C.; LEARDI, R.; DRAVA, G. A class modeling technique based on potential functions. *Journal of Chemometrics*, v. 5, p. 435-453, 1991. <https://onlinelibrary.wiley.com/doi/abs/10.1002/cem.1180050504>
- GAITÁN-JURADO, A. J.; GARCÍA-MOLINA, M.; PEÑA-RODRÍGUEZ, F.; ORTIZ-SOMOVILLA, V. Near infrared applications in the quality control of seed cotton. *Journal of Near Infrared Spectroscopy*, v. 16, n. 4, p. 421-429, 2008. <https://journals.sagepub.com/doi/abs/10.1255/jnirs.806>
- GAMA, J. S. N.; BRUNO, R. L. A.; SILVA, K. R. G.; RÊGO, E. R.; PEREIRA FILHO, T. B.; BARBOSA, R. C.; BEZERRA, A. K. D. Qualidade fisiológica e sanitária de sementes de erva-doce (*Foeniculum vulgare* Mill.) armazenadas. *Revista Brasileira de Plantas Mediciniais*, v.14, n.esp., p.175-182, 2012. http://www.scielo.br/scielo.php?pid=S1516-05722012000500009&script=sci_arttext
- GRUNVALD, A. K.; CARVALHO, C. G. P. D.; LEITE, R. S.; MANDARINO, J. M. G.; ANDRADE, C. A. D. B.; SCAPIM, C. A. Predicting the oil contents in sunflower genotype seeds using near-infrared reflectance (NIR) spectroscopy. *Acta Scientiarum: Agronomy*, v. 36, n. 2, p. 233-237, 2014. http://www.scielo.br/scielo.php?pid=S1807-86212014000200013&script=sci_arttext
- 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*, v. 43, p. 654-660, 2013. <https://www.sciencedirect.com/science/article/abs/pii/S0926669012004852>
- ISHIKAWA, H.; BOUKAR, O.; FATOKUN, C.; SHONO, M.; MURANAKA, S. Development of calibration model to predict nitrogen content in single seeds of cowpea (*Vigna unguiculata*) using near infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, v. 25, n. 3, p. 211-214, 2017. <https://journals.sagepub.com/doi/abs/10.1177/0967033517712129>
- JIA, S.; AN, D.; LIU, Z.; GU, J.; LI, S.; ZHANG, X.; ZHU, D.; GUO, T.; YAN, Y. et al. Variety identification method of coated maize seeds based on near-infrared spectroscopy and chemometrics. *Journal of Cereal Science*, v. 63, p. 21-26, 2015. <https://www.sciencedirect.com/science/article/abs/pii/S073352101400126X>
- MAGUIRE, J. D. Speed of germination – aid in selection and evaluation for seedling emergence and vigor. *Crop Science*, v. 2, n. 2, p. 176-177, 1962. <https://acsess.onlinelibrary.wiley.com/doi/abs/10.2135/cropsci1962.0011183X000200020033x>
- MATTIONI, F.; ALBUQUERQUE, M.C.F.; GUIMARÃES, S.C. Vigor de sementes e desempenho agrônômico de plantas de algodão. *Journal of Seed Science*, v. 34, n. 1, p. 108-116, 2012. http://www.scielo.br/scielo.php?pid=S0101-31222012000100014&script=sci_arttext
- NUNES, C. A.; FREITAS, M. P.; PINHEIRO, A. C. M.; BASTOS, S. C. Chemoface: a novel free user-friendly interface for chemometrics. *Journal of the Brazilian Chemical Society*, v. 23, n. 11, p. 2003-2010, 2012. http://www.scielo.br/scielo.php?pid=S0103-50532012005000073&script=sci_arttext
- PANDE, R.; MISHRA, H. N. Fourier transform near-infrared spectroscopy for rapid and simple determination of phytic acid content in green gram seeds (*Vigna radiata*). *Food Chemistry*, v. 172, p. 880-884, 2015. <https://www.sciencedirect.com/science/article/pii/S0308814614014162>
- PEDROSO, D. C.; MENEZES, V. O.; MUNIZ, M. F. B.; PIVETA, G.; TUNES, L. M. D.; MULLER, J.; MENEZES, N. L. D.. Métodos de inoculação de *Alternaria alternata* e *A. dauci* em sementes de salsa e sua influência na qualidade fisiológica. *Journal of Seed Science*, v.32, n.3, p.79-85, 2010. http://www.scielo.br/scielo.php?pid=S0101-31222010000300009&script=sci_arttext&tlng=pt

- QUEIROGA, V. P.; CASTRO, L. B. Q.; PALMEIRA, J.; GOMES, A. L. D. S.; ALVES, N. M. C.; ARAÚJO, D. R. Qualidade fisiológica de sementes de algodão armazenadas em função de diferentes tratamentos e cultivares. *Revista Brasileira de Produtos Agroindustriais*, v.11, n.1, p.43-54, 2009. <http://www.deag.ufcg.edu.br/rbpa/rev111/Art1117.pdf>
- RINGSTED, T.; DUPONT, S.; RAMSAY, J.; JESPERSEN, B. M.; SØRENSEN, K. M.; KEIDING, S. R.; ENGELSEN, S. B. Near-infrared spectroscopy using a supercontinuum laser: application to long wavelength transmission spectra of barley endosperm and oil. *Applied Spectroscopy*, v. 70, n. 7, p. 1176-1185, 2016. <https://journals.sagepub.com/doi/abs/10.1177/0003702816652361>
- ROSSATO, R.; PRETE, C. E. C.; CASTRO, C.; TOMM, G. O.; LEITE, R. S.; MANDARINO, J. M. G.; ARAÚJO, P. M.; CARVALHO, C. G. P. Predicting rapeseed oil content with near-infrared spectroscopy. *Pesquisa Agropecuária Brasileira*, v. 48, n. 12, p. 1601-1605, 2013. http://www.scielo.br/scielo.php?pid=S0100-204X2013001200010&script=sci_arttext
- SANTOS, M. B. H.; GOMES, A. A.; VILAR, W. T. S.; ALMEIDA, P. B. A.; MILANI, M.; NÓBREGA, M. B. M.; MEDEIROS, E. P.; GALVÃO, R. K. H.; ARAÚJO, M. C. U. Non-destructive NIR spectrometric cultivar discrimination of Castor seeds resulting from Breeding Programs. *Journal of the Brazilian Chemical Society*, v. 25, n. 5, p.969-974, 2014. http://www.scielo.br/scielo.php?pid=S0103-50532014000500021&script=sci_arttext
- SHETTY, N.; OLESEN, M. H.; GISLUM, R.; DELEURAN, L. C.; BOELT, B. Use of partial least squares discriminant analysis on visible-near infrared multispectral image data to examine germination ability and germ length in spinach seeds. *Journal of Chemometrics*, v. 26, n. 8-9, p. 462-466, 2012. <https://onlinelibrary.wiley.com/doi/abs/10.1002/cem.1415>
- 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*, v. 237, p. 1027-1034, 2016. <https://www.sciencedirect.com/science/article/abs/pii/S0925400516314034>
- SOARES, S. F. C.; MEDEIROS, E. P.; PASQUINI, C.; DE LELIS 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. <https://pubs.rsc.org/en/content/articlelanding/2016/ay/c6ay02896a/unauth#!divAbstract>
- STEINER, F.; OLIVEIRA, S. S. C. D.; MARTINS, C. C.; CRUZ, S. J. S. Comparação entre métodos para a avaliação do vigor de lotes de sementes de triticale. *Ciência Rural*, v. 41, n. 2, p. 200-204, 2011. http://www.scielo.br/scielo.php?pid=S0103-84782011000200002&script=sci_arttext&tlng=pt
- SZYMAŃSKA, E.; SACCENTI, E.; SMILDE, A. K.; WESTERHUIS, J. A. Double-check: validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Springer US*, 2011. <https://link.springer.com/article/10.1007/s11306-011-0330-3>
- VIEIRA, M. G. G. C.; VON PINHO, E. V. R. Metodologia do teste de tetrazólio em sementes de algodão. In: KRZYŻANOWSKI, F. C.; VIEIRA, R. D.; FRANÇA- NETO, J. B. (Eds.). *Vigor de sementes: conceitos e testes*. Londrina: ABRATES. Cap. 8, p.8.1-8.13. 1999.
- ZORZAL, T. A.; FANTINATO, D. E.; CAMPOS, L. M.; LUZ, A. C. C.; CORTE, V. B. Teste do tetrazólio para estimativa da viabilidade de sementes. *Natureza on line*, v. 13, n. 3, p.144-149, 2015. http://www.naturezaonline.com.br/natureza/conteudo/pdf/ZorzalTAetal_144-149.pdf

