



Seed quality of *Brachiaria brizantha* by X-ray image analysis using a Bayesian approach

Alessandra Querino da Silva^{1*}, Luciano Antonio de Oliveira¹, Carlos Pereira da Silva², Cristian Tiago Erazo Mendes², Ana Maria Oliveira Ferreira², Thelma Sáfiadi² and Maria Laene Moreira de Carvalho²

¹Faculdade de Ciências Exatas e Tecnologia, Universidade Federal da Grande Dourados, Rodovia Dourados-Itahum, Km 12, Cx. Postal 364, 79804-970, Dourados, Mato Grosso do Sul, Brazil. ²Universidade Federal de Lavras, Lavras, Minas Gerais, Brazil. *Author for correspondence. E-mail: alessandrasilva@ufgd.edu.br

ABSTRACT. *Brachiaria brizantha* is the most economically important forage grass in Brazil and is propagated through sowing. Producing high-quality seeds has been a constant challenge due to their uneven maturation. The development and application of precise and non-destructive methods for identifying internal damages to seeds, such as the X-ray test, which quickly indicates the quality of the lots, is of fundamental importance for the seed industry. In this work, the quality of *Brachiaria brizantha* seeds was analyzed based on the morphological characteristics observed in X-ray images that were related to viability using a mixture model under a Bayesian approach, with the following objectives: i) verify the adequacy of the Bayesian modeling used in the data analysis; ii) associate the efficiency of using radiographs as a way to assess the viability of the seeds; and iii) relate the classifications carried out by evaluators with the probability of originating normal or abnormal seedlings. The methodology applied for the analysis proved to be adequate. Further, the Bayesian estimates for parameters related to internal morphology were established with associated levels of uncertainty, which represents an advantage over usual frequentist methods. Based on the model's estimates, seeds evaluated as potentially unviable by three evaluators had practically no probability of germination and did not germinate in the test applied later. Seeds classified as potentially viable had a high probability of developing into normal seedlings, while 73.27% showed this property in the germination test.

Keywords: mixture model; Bayesian inference; X-ray test.

Received on September 23, 2020.

Accepted on January 17, 2021.

Introduction

Pasture is the preferred form for forage production, which constitutes the natural basis of feeding herbivores kept in field conditions. Brazil stands out in the world in terms of its production, consumption, and export of forage grass seeds, with *Brachiaria brizantha* being the most important in terms of commercialization (Pereira, Oliveira, Rosa, & Kikuti, 2011; Silva, Martins, Cruz, Jeromini, & Bruno, 2017). Due to the tolerance and adaptability of *Brachiaria* to Brazilian soil and climatic conditions, as well as the relative ease of its management, it became responsible for advancing national livestock. Additionally, it provides pastures of reasonable quality and moderate investments.

The use of high-quality seeds is essential because the establishment of pasture occurs through sowing (Batista, Cardoso, Binotti, Costa, & Sá, 2016a; Batista, Binotti, Cardoso, Costa, & Nascimento, 2016b). Thus, the need to obtain them has motivated the development of increasingly efficient production systems, as well as analysis procedures and methods that support the selection of seeds for cultivation (Cicero, Van Der Heijden, Van Der Burg, & Bino, 1998). Seed quality involves different attributes including viability. However, information related to viability, among other aspects, is difficult to obtain due to restrictions related to time and technological aspects (Medeiros, Pinheiro, Xavier, Silva, & Dias, 2020).

Regarding *Brachiaria brizantha* seeds, the focus of this work, the long period of time in germination tests combined with seed dormancy contributes to the tetrazolium test being routinely used to estimate viability (Novembre, Chamma, & Gomes, 2006; Cardoso et al., 2014; Batista et al., 2016a). It is a biochemical test that demands time and expertise of the analysts.

An interesting alternative for seed quality tests is the study of the internal morphology of seeds through the analysis of X-ray images. The use of X-ray images in seed evaluation was pioneered by Simak and

Gustafsson (1953) and has since found wide applicability. Methodologies have been proposed to improve seed quality with respect to its physical and physiological aspects, providing a rapid assessment of most species, including forage grasses, such as *Brachiaria* (Gomes Junior, 2010; Lima, Santos, Dutra, & Barros, 2013; Gomes, Martins, & Gomes Junior, 2014; Jeromini, Martins, Pereira, & Gomes Junior, 2019).

X-rays do not affect germination, making it a quick, simple, and nondestructive test to perform. Other than identifying morphological aspects possibly associated with germination viability, they also allow an additional posterior analysis of the same seed (Copeland & McDonald, 2001; Bino, Aartse, & Van Der Burg, 1993; Menezes, Cícero, & Villela, 2005). Furthermore, methodological advances using X-rays allow the analysis of immature or aberrated embryos, predicting the presence of abnormal seedlings (Mondo & Cicero, 2005).

The association of statistical methodologies for analyzing data from radiographic images is also of fundamental importance. In this context, the use of mixture models offers great versatility for analysis; when the mechanism that generates the observed data is unknown, they can be applied in various contexts (Marin, Mengersen, & Robert, 2005; Rufo, Martín, & Pérez, 2006; Schlattmann, 2009). For this purpose, it assumes that n elements in the sample come from J distinct and unidentified subgroups (or populations) in different proportions and, generally, the objective is to analyze these different populations and/or classify individuals according to their group of origin (Diebolt & Robert, 1994).

Several methods have been proposed for the estimation of parameters in mixture models. When own priorities are available, the Bayesian method can be used and the analysis is simplified by introducing latent variables with a hierarchical structure (Tanner & Wong, 1987).

Procedures belonging to the class of Markov Chain Monte Carlo (MCMC) method, especially the Gibbs sampler, allow sampling values from a posteriori conditional densities in complex multivariate models and represent a great advantage of the Bayesian inference (Simpson, Rue, Riebler, Martins, & Sorbye, 2017). Bayesian analyses have been applied in research studies associated with seed quality (Junqueira et al., 2016) and, especially, using a mixture model to evaluate the viability potential (Rocha, Sáfadi, & Carvalho, 2013).

This work sought to associate morphological aspects to the quality of *B. brizantha* seeds through the visual inspection of X-ray images performed by three evaluators and to relate them with viability. The problem was modeled by a mixture of binomial distributions, and the parameters were estimated using the Bayesian method, with the following objectives: i) verify the adequacy of the Bayesian modeling used in the data analysis; ii) associate the efficiency of using radiographs as a way to assess the viability of the seeds; and iii) relate the classifications carried out by evaluators with the probability of originating normal or abnormal seedlings.

Material and methods

Experiment

Four hundred seeds were randomly selected from the seed lot of *B. brizantha* cv. Marandu produced in the state of Mato Grosso in the 2016/2017 harvest. The experiment was conducted at the Central Seed Laboratory of the Department of Agriculture of the Federal University of Lavras (UFLA) in Lavras, Minas Gerais State, Brazil, according to the following procedure:

1) Obtaining radiographic images of the seeds: The seeds were randomly subdivided into 8 subsamples of 50 seeds arranged and fixed on an overhead transparency sheet using double-sided tape. They were numbered according to the rows and columns they occupied for identification purposes. They were then radiographed, without any preparation, with an intensity of 35 kV and an average of 14 seconds of exposure in a Faxitron HP Model MX20 device to generate digital images (Figure 1).

The images obtained from each seed were individually analyzed by three specialists and classified as Full, Partially Full, Damaged, or Empty (Figure 2). The seeds classified as empty or damaged were considered potentially unviable.

After the X-ray images were obtained, the seeds underwent the germination test according to the Rules for Seed Analysis (RAS) (Brasil, 2009).

2) Germination test: This test was conducted by sowing the seeds onto two sheets of white blotting paper moistened with distilled water equivalent to 2.5 times the dry mass of the papers. They were arranged in acrylic boxes like gerbox (Figure 3), and maintained in a Biochemical Oxygen Demand (BOD) chamber with alternating temperatures of 20 - 35°C under a photoperiod of 12 hours of light and 12 hours of darkness.

The evaluations of normal seedlings were performed according to the criteria established by RAS on the 7th, 14th, and 21st day after sowing.

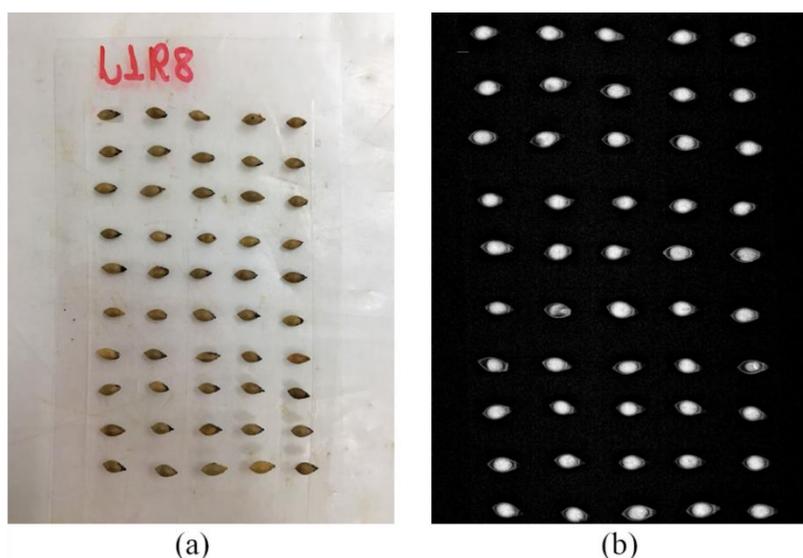


Figure 1. (a) seeds arranged for X-ray testing and (b) scanned image.

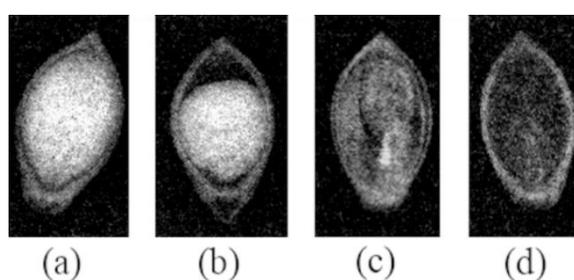


Figure 2. X-ray images of *Brachiaria brizantha* seeds classified into four categories: Full (a), Partially Full (b), Damaged (c), or Empty (d).

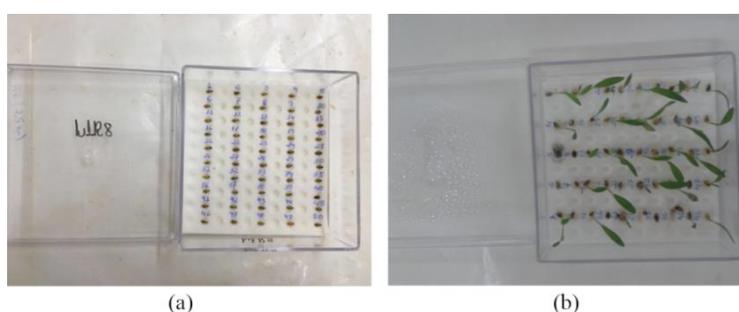


Figure 3. (a) Preparation of seeds for germination and (b) Germination at seven days.

After completing the germination test, the remaining seeds were subjected to the tetrazolium test to verify their viability.

3) Tetrazolium test: The remaining seeds were cut longitudinally using tweezers and a scalpel. They were then immersed in a 0.5% solution of 2-3-5 triphenyl tetrazolium chloride in dark flasks and kept in a BOD-type germination chamber at 30°C for three hours (Brasil, 2009). After which, the seeds were evaluated as viable or non-viable based on the location and intensity of the color of their structures.

Relative frequency method

Each of the 400 radiographic images of seeds was classified as potentially unviable by three independent experts. We defined the number of evaluations of X-ray images of potentially unviable ($Y = 0, 1, 2,$ and 3) *Brachiaria* seeds as random variable Y . The possible results are as follows:

- 0: None of the three evaluators classified the radiographic image of the *Brachiaria* seed as potentially unviable;
- 1: One of the three evaluators classified the images as potentially unviable;
- 2: Two of the three evaluators classified the image as potentially unviable;
- 3: The three evaluators classified the image as potentially unviable.

Let event N : the seed developed into a normal seedling and \bar{N} : the seed either developed into an abnormal seedling or did not germinate. We also consider that:

λ = probability that the seed will develop into a normal seedling ($P(N)$);

p_1 = probability that the seed will develop into a normal seedling after being evaluated as potentially unviable; and

p_2 = probability that the seed will either germinate an abnormal seedling or not germinate, since it has been assessed as potentially unviable.

The conditional distributions of the variable Y given the occurrence of events N and \bar{N} : are independent and given by binomials with parameters $(3, p_1)$ and $(3, p_2)$, respectively:

$$P(Y = y|N) = C_{3,y}p_1^y(1-p_1)^{3-y} \text{ and } P(Y = y|\bar{N}) = C_{3,y}p_2^y(1-p_2)^{3-y} \quad (1)$$

In turn, the marginal distribution of Y is:

$$\begin{aligned} P(Y = y) &= P(N)P(Y = y|N) + P(\bar{N})P(Y = y|\bar{N}) = \\ &= \lambda C_{3,y}p_1^y(1-p_1)^{3-y} + (1-\lambda)C_{3,y}p_2^y(1-p_2)^{3-y} \text{ where } C_{3,y} = \binom{3}{y} \end{aligned} \quad (2)$$

To estimate the parameters λ , p_1 , and p_2 in equation (2), it is necessary to solve the system given by:

$$\begin{cases} P(Y = 0) = \lambda(1-p_1)^3 + (1-\lambda)(1-p_2)^3 \\ P(Y = 1) = 3\lambda p_1(1-p_1)^2 + 3(1-\lambda)p_2(1-p_2)^2 \\ P(Y = 2) = 3\lambda p_1^2(1-p_1) + 3(1-\lambda)p_2^2(1-p_2) \\ P(Y = 3) = \lambda p_1^3 + (1-\lambda)p_2^3 \end{cases} \quad (3)$$

The solution of equation (3) was obtained using the Newton–Raphson iterative method (Ruggiero & Lopes, 1997).

According to Bayes' theorem, the probability of the seed developing into normal seedlings given the number of evaluations as potentially unviable is:

$$\begin{aligned} P(N|Y = y) &= \frac{P(Y = y|N) \cdot P(N)}{P(Y = y)} = \frac{\lambda C_{3,y}p_1^y(1-p_1)^{3-y}}{\lambda C_{3,y}p_1^y(1-p_1)^{3-y} + (1-\lambda)C_{3,y}p_2^y(1-p_2)^{3-y}} \\ \Rightarrow P(N|Y = y) &= \frac{\lambda p_1^y(1-p_1)^{3-y}}{\lambda p_1^y(1-p_1)^{3-y} + (1-\lambda)p_2^y(1-p_2)^{3-y}} \text{ with } y = 0,1,2,3. \end{aligned} \quad (4)$$

From these estimators, it is possible to establish the expected number of seeds that developed into normal seedlings ($E(N_i)$) depending on the number of classifications as potentially unviable ($Y = y$). $E(N_i)$ is obtained by the product of $P(N|Y = y)$ with the observed number of images classified as potentially unviable in each category.

Mixture model of two binomial distributions

Expression (2) shows that the problem addressed can be modeled by a mixture of distributions (Stephens, 2000).

Considering the binomial components, the density of Y is expressed as:

$$p(y|\Psi) = \lambda \binom{3}{y} p_1^y(1-p_1)^{3-y} + (1-\lambda) \binom{3}{y} p_2^y(1-p_2)^{3-y} \quad (5)$$

where $\Psi = (\lambda, p_1, p_2)$, λ , p_1 and p_2 are quantities that take on values in the range $[0,1]$. The likelihood function, for $n = 400$, is given by

$$L(\Psi|y) = \prod_{i=1}^n \left(\lambda \binom{3}{y_i} p_1^{y_i}(1-p_1)^{3-y_i} + (1-\lambda) \binom{3}{y_i} p_2^{y_i}(1-p_2)^{3-y_i} \right) \quad (6)$$

After assuming a model for the data, the next step is to assign *a priori* information to the unknown quantities.

Priors distributions

The *a priori* densities for parameters λ and p_j ($j = 1, 2$) are, respectively,

$$\lambda \sim \text{Beta}(a_\lambda, b_\lambda) = \frac{\Gamma(a_\lambda+b_\lambda)}{\Gamma(a_\lambda)\Gamma(b_\lambda)} \lambda^{a_\lambda-1}(1-\lambda)^{b_\lambda-1} \quad (7)$$

$$p_j \sim \text{Beta}(a_{p_j}, b_{p_j}) = \frac{\Gamma(a_{p_j}+b_{p_j})}{\Gamma(a_{p_j})\Gamma(b_{p_j})} p_j^{a_{p_j}-1}(1-p_j)^{b_{p_j}-1} \quad (8)$$

where $\Gamma(\cdot)$ denotes a Gamma function.

Equation (6) can be simplified by introducing latent random variables, which indicate the subpopulation each observation that make up \mathbf{y} belongs to, by using the augmented data structure (Tanner & Wong, 1987). For each y_i observation, it is always possible to associate a latent allocation variable Z_i in a way that $Z_i = j$ indicates that y_i belongs to the j -th component of the mixture. Thus, Z_i has a Bernoulli distribution denoted by $Z_i \sim \text{Bernoulli}(\lambda)$.

As already pointed out, the latent variables introduced were intended to classify each observation $Y = y_{ij}$ in one of two populations with $j = 1, 2$. The term n_j indicates the number of observations classified in each subpopulation, where $n_j = \sum_{i=1}^n \sum_{j=1}^2 z_{ij} I(y_i \in A_j)$ (if $z_{ij} = 1$ then $y_i \in A_j$, and A_j a subpopulation; otherwise $z_{ij} = 0$). The joint likelihood is:

$$L(\Psi|y, z) = \lambda^{n_1} \left[\prod_{i=1}^{n_1} \binom{3}{y_{i1}} \right] p_1^{\sum_{i=1}^{n_1} y_{i1}} (1 - p_1)^{3n_1 - \sum_{i=1}^{n_1} y_{i1}} \times \\ \times (1 - \lambda)^{n_2} \left[\prod_{i=1}^{n_2} \binom{3}{y_{i2}} \right] p_2^{\sum_{i=1}^{n_2} y_{i2}} (1 - p_2)^{3n_2 - \sum_{i=1}^{n_2} y_{i2}} \quad (9)$$

Complete posterior conditional densities

The posterior joint distribution was obtained by combining the likelihood function and prior densities using Bayes' theorem:

$$p(\Psi|y, z) \propto \lambda^{n_1} p_1^{\sum_{i=1}^{n_1} y_{i1}} (1 - p_1)^{3n_1 - \sum_{i=1}^{n_1} y_{i1}} (1 - \lambda)^{n_2} p_2^{\sum_{i=1}^{n_2} y_{i2}} (1 - p_2)^{3n_2 - \sum_{i=1}^{n_2} y_{i2}} \times \\ \times \lambda^{a_\lambda - 1} (1 - \lambda)^{b_\lambda - 1} \times \prod_{j=1}^2 p_j^{a_{p_j} - 1} (1 - p_j)^{b_{p_j} - 1} \quad (10)$$

Conditional distributions were obtained using algebraic manipulations of equation (10). The complete posterior conditional distribution for p_j is given by:

$$p(p_j | \dots) \propto p_j^{\sum_{i=1}^{n_j} y_{ij}} (1 - p_j)^{3n_j - \sum_{i=1}^{n_j} y_{ij}} \times p_j^{a_{p_j} - 1} (1 - p_j)^{b_{p_j} - 1} \\ \propto p_j^{\sum_{i=1}^{n_j} y_{ij} + a_{p_j} - 1} (1 - p_j)^{3n_j - \sum_{i=1}^{n_j} y_{ij} + b_{p_j} - 1} \quad (11)$$

therefore $p_j | \dots \sim \text{Beta} \left(\sum_{i=1}^{n_j} y_{ij} + a_{p_j}, 3n_j - \sum_{i=1}^{n_j} y_{ij} + b_{p_j} \right)$ with $j = 1, 2$.

Similarly, we find the complete conditional distribution for the weight λ as:

$$p(\lambda | \dots) \propto \lambda^{n_1} (1 - \lambda)^{n_2} \lambda^{a_\lambda - 1} (1 - \lambda)^{b_\lambda - 1} \\ \propto \lambda^{n_1 + a_\lambda - 1} (1 - \lambda)^{n_2 + b_\lambda - 1}$$

which is also a beta distribution, $\lambda | \dots \sim \text{Beta}(n_1 + a_\lambda, n_2 + b_\lambda)$. Finally, there is a *posteriori* conditional distribution for z_i is:

$$p(Z = z_i | \dots) \propto \prod_{j=1}^k \lambda_j^{n_j} p_j^{y_{ij}} (1 - p_j)^{3 - y_{ij}} p_j^{a_{p_j} - 1} (1 - p_j)^{b_{p_j} - 1} \lambda^{a_\lambda - 1} (1 - \lambda)^{b_\lambda - 1}.$$

Considering $a_\lambda = b_\lambda = a_{p_j} = b_{p_j} = 1$, it is possible to obtain non-informative (uniform) prior densities.

The augmented data method is flexible for a Bayesian hierarchical treatment sampling of indicator variables Z_i (Bilancia & Pollice, 1999). The posterior probability of allocation for y_i in the population $j = 1$, $w_{i1}(z_i = 1)$, is given by:

$$w_{i1} | \lambda, p_1, p_2 = \frac{\lambda p(y_i | p_1)}{\lambda p(y_i | p_1) + (1 - \lambda) p(y_i | p_2)}, i = 1, \dots, n \\ w_{i1} | \lambda, p_1, p_2 = \frac{\lambda \binom{3}{y_i} p_1^{y_i} (1 - p_1)^{3 - y_i}}{\lambda \binom{3}{y_i} p_1^{y_i} (1 - p_1)^{3 - y_i} + (1 - \lambda) \binom{3}{y_i} p_2^{y_i} (1 - p_2)^{3 - y_i}}, i = 1, \dots, n.$$

Consequently, Z_i will be sampled from a Bernoulli distribution with w_{i1} parameter.

Sampling process

The sampling process was carried out using the Markov Chain Monte Carlo (MCMC) method. It is an iterative stochastic process based on the successive simulation of values from a *posteriori* complete

conditional density. Because the conditional densities have known shapes and allow direct sampling, a Gibbs sampler was used. Its algorithm is given by:

- i) Assign initial values $\Psi^0 = (\lambda^0, p_1^0, p_2^0)$
- ii) For iteration t :
 - a) Calculate $w_{i1}^t | \lambda^{t-1}, p_1^{t-1}, p_2^{t-1}$ with $i=1, \dots, n$;
 - b) Sample z_i^t from Bernoulli (w_{i1}^t) for i from 1 to n ;
 - c) Sample $\lambda^t | n_1^t, n_2^t, z^t$;
 - d) Sample $p_1^t | n_1^t, S_1^t, z^t$ with $S_1^t = \sum_{i=1}^{n_1^t} y_{i1}$;
 - e) Sample $p_2^t | n_2^t, S_2^t, z^t$ with $S_2^t = \sum_{i=1}^{n_2^t} y_{i2}$.

The convergence of the chains was monitored by Raftery and Lewis criteria (Raftery & Lewis, 1992) and Heidelberger and Welch criteria (Heidelberger & Welch, 1983), implemented in the "boa" package (Smith, 2007) in R (R Core Team, 2017).

The expressions obtained for the complete *a posteriori* conditional densities, as well as those presented for the method of relative frequency, are the same in Achcar and Ruffino Netto (2003) and Rocha et al. (2013).

Point estimates were obtained by the posterior mean of the simulated MCMC chains, and interval estimates were implemented through the highest posterior density (HPD) regions at 95% credibility.

The analyses were performed using resources in the "gtools" package (Warnes, Bolker, & Thomas, 2018) in software R.

Results

The frequencies referring to the number of seeds classified as potentially unviable by the three evaluators are shown in Table 1. It was observed that 303 seeds were evaluated as potentially viable, which corresponds to 75.75% of the total seeds. On the other hand, 13.5% of them (54 seeds) were considered potentially unviable. For the remaining 43 seeds, there was no consensus between the evaluators' ratings of the images.

Table 1. Frequency distribution of the number of radiographic images of *Brachiaria brizantha* seeds classified as potentially unviable by three evaluators.

| Y: Number of classifications as potentially unviable | Number of radiographic images (f) |
|--|-----------------------------------|
| 0 | 303 |
| 1 | 30 |
| 2 | 13 |
| 3 | 54 |
| Total | 400 |

Markov chains were simulated with 41,634 iterations using the Gibbs sampler algorithm. The first 12 observations were discarded as a warm-up period ("*burn-in*"). To obtain an uncorrelated sample, a spacing (jump) between the sampled points of three iterations ("*thinning*") was considered, resulting in a final sample of size 13,866. These burn-in and jump values were based on a training sample according to the Raftery and Lewis test (1992).

Table 2 presents the results obtained from the convergence tests using the methods of Raftery and Lewis (1992) and Heidelberger and Welch (1983). It was observed that all parameters had a dependency factor of $I < 5$ (Raftery & Lewis, 1992). In addition, the chains for all parameters passed on the stationarity test, indicating that convergence was also achieved according to the criteria of Heidelberger and Welch (1983).

Table 2. Results of tests to monitor convergence of Markov chains.

| Parameter | Raftery and Lewis I | Heidelberger and Welch Stationarity test | Test Halfwidth |
|-----------|------------------------|---|----------------|
| λ | 1.01548 | Yes | Yes |
| p_1 | 1.03364 | Yes | Yes |
| p_2 | 1.01548 | Yes | Yes |

I: dependency factor.

Figure 4 shows the trace plots and density plots of the Markov chain samples for each parameter. It is observed that the values oscillate around a constant value, which is indicative of stationarity. Thus, the

properties visualized in these representations corroborate the results of the tests used to monitor the convergence of the chains. There is also a greater symmetry for the densities of λ and p_1 as well as a slight asymmetry to the left with respect to the density of p_2 .

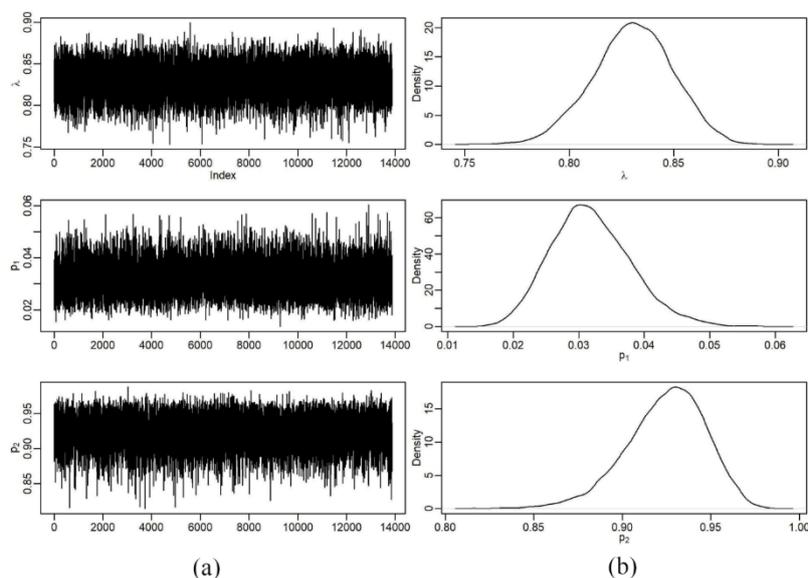


Figure 4. (a) Traces of the MCMC chains and (b) graphical representation of *a posteriori* densities for parameters λ , p_1 and p_2 .

Table 3 shows posterior means, standard deviations (SD), and HPD intervals at 95% credibility for the parameters of model λ , p_1 , and p_2 . As previously emphasized, λ assesses the probability that the seed develops into a normal seedling with *a posteriori* mean of 0.8304. The probability of the seed developing into an abnormal seedling or not germinating (0.1696) is referred to as $(1 - \lambda)$.

Table 3. Posterior mean, standard deviation (SD), and HPD interval at 95% credibility (LL: lower limit, UL: upper limit) for parameters λ , p_1 , and p_2 .

| Parameter | Mean | SD | HPD interval | |
|-----------|--------|--------|--------------|--------|
| | | | LL | UL |
| λ | 0.8304 | 0.0192 | 0.7925 | 0.8670 |
| p_1 | 0.0319 | 0.0061 | 0.0200 | 0.0437 |
| p_2 | 0.9244 | 0.0226 | 0.8806 | 0.9668 |

On the other hand, it can be seen from the estimates that the conditional probability of the seed being evaluated as potentially unviable is small among those that developed into normal seedlings. It has posterior mean of 0.0319 and a small standard deviation (0.0061), which shows coherence among the evaluators regarding the physiological aspects observed from the images. The probability that the seed was classified as potentially unviable among those that did not germinate (or developed into abnormal seedlings) was high (0.9244).

Table 4 shows means, SDs, and regions of HPD credibility for the probability of the seed having developed into a normal seedling conditional on the realization of each value of the random variable $Y = y$ ($y = 0, 1, 2, 3$), that is, the number of evaluations as potentially unviable. For a seed that has not received any evaluation as potentially unviable ($Y = 0$), the event developing into a normal seedling is practically certain, since the posterior mean of the conditional probability is very close to one (0.9999) and the SD is approximately equal to zero (0.0001). There is also a high probability of developing into normal seedlings conditioned to $Y = 1$. For seeds whose number of evaluations as unviable was two or three, the conditional probabilities were low. Point and interval estimates were based on the combined MCMC samples to obtain other variables that were not initially considered in the model.

Posterior point and interval estimates for the expected number of normal seedlings in each of the classifications are shown in Table 5. These values were obtained by the product between the frequencies observed in each class and the conditional probability of normal seedlings given the occurrence of the number of X-ray images considered potentially unviable.

Table 4. Posterior mean, standard deviation (SD), and lower (LL) and upper (UL) limits of HPD regions at 95% credibility regarding the conditional probability of the seed to develop into normal seedling given the number of evaluations as potentially unviable.

| Y | Mean | SD | HPD interval | |
|---|--------|--------|--------------|--------|
| | | | LL | UL |
| 0 | 0.9999 | 0.0001 | 0.9996 | 1.0000 |
| 1 | 0.9607 | 0.0258 | 0.9095 | 0.9967 |
| 2 | 0.0789 | 0.0412 | 0.0154 | 0.1567 |
| 3 | 0.0002 | 0.0001 | 0.0000 | 0.0005 |

Table 5. Posterior mean, standard deviation (SD), and lower (LL) and upper (UL) limits of HPD regions at 95% credibility for the expected number of normal seedlings in each of the image classifications as potentially unviable.

| Y | Mean | SD | HPD interval | |
|---|----------|--------|--------------|----------|
| | | | LL | UL |
| 0 | 302.9615 | 0.0389 | 302.8895 | 302.9996 |
| 1 | 28.8202 | 0.7737 | 27.2856 | 29.9001 |
| 2 | 1.0255 | 0.5351 | 0.2006 | 2.0365 |
| 3 | 0.0123 | 0.0077 | 0.0018 | 0.0274 |

For comparison, estimates are presented for λ , p_1 and p_2 obtained by the method of relative frequencies, usually used in these circumstances (Achcar & Ruffino Netto, 2003; Rocha et al., 2013). The parameter estimates were: $\lambda = 0.8325$, $p_1 = 0.03102$, and $p_2 = 0.9305$.

From the relative frequency method, the probability of obtaining normal seedlings, conditional on the realized value of the random variable, it is found by the solution of equation (4). In addition, from the data presented in Table 1 and estimates obtained by the method of relative frequency, calculating the expected number of normal seedlings is possible. These results are shown in Table 6.

Table 6. Frequency distribution of the evaluation of 400 *Brachiaria* seed images by three evaluators, and conditional probabilities and expected number of normal seedlings using the relative frequency method.

| Number of classifications as potentially unviable (Y) | Number of seed images (f_i) | $P(N Y=y)$ | $E(N_i) = P(N Y=y)(f_i)$ |
|---|---------------------------------|------------|--------------------------|
| 0 | 303 | 0.9999 | 302.9776 |
| 1 | 30 | 0.9699 | 29.0982 |
| 2 | 13 | 0.0716 | 0.9305 |
| 3 | 54 | 0.0002 | 0.0099 |

It is possible to observe that the Bayesian point solutions are very close to those obtained through the method of relative frequency. Both methods provided good estimates, since the expected number of normal seedlings as a function of $Y = 0$ was approximately 303, equal to the observed number of images classified by the three evaluators as potentially viable. In addition, when the random variable takes the value $Y = 3$, the expected number of normal seedlings is very small (0.0099), which is in accordance with the probability of this event occurring (0.0002).

In the germination test of 400 seeds, 240 developed into normal seedlings, 80 did not germinate, and the others developed into abnormal seedlings. The remaining seeds were subjected to the tetrazolium test, and only seven were identified as viable. Of the 54 seeds classified as potentially unviable by the three evaluators, none germinated and were considered unviable in the tetrazolium test. It was also observed that most of the seeds classified as potentially viable ($Y = 0$) developed into normal seedlings.

Discussion

The germination test was completely available for 21 days, which is a long time to decide on a production system (Gaspar-Oliveira, Martins, Nakagawa, & Cavariani, 2008). On the other hand, obtaining radiographic images, classifying the viability of the seeds, and the statistical analysis based on the mixture model took up less time. In this sense, a X-ray test for *Brachiaria brizantha* cv. Marandu is a supporting procedure, especially because it is already indicated by the RAS. In addition, the results highlight the advantages of using

noninvasive optical technologies that evaluate the germination of the same seeds submitted to X-ray tests, among others, since this method is nondestructive. However, the X-ray test does not replace the germination test, which defines the potential of a seed lot to produce normal seedlings under favorable conditions.

Of the total seeds that were classified as potentially viable ($Y = 0$) by the three evaluators, 222 developed into normal seedlings, corresponding to 73.27%, which is statistically lower than estimated by the Bayesian model, being outside the region of credibility for the conditional probability $P(N|Y = 0)$ (Table 5). It is necessary to emphasize that not all seeds classified as viable will develop into normal seedlings. This highlights the importance of performing the germination test.

In the evaluated lot, 60% germinated, which is the minimum for the commercialization of *B. brizantha* (Brasil, 2008). A batch of seeds with low purity was used with the intention that different types of classifications (filled, partially filled, deformed, and empty) would occur, especially to allow evaluators to use the knowledge they obtained in the previous training. Furthermore, it was used for the method to be applied in a more illustrative way.

Of the seeds classified as potentially unviable by the three evaluators, none germinated and were also considered unviable by the results of the tetrazolium test. This indicates an efficient association between morphological attributes and the non-viability of germination. This result is compatible with the model estimate.

The method applied in this work was also used by Rocha et al. (2013) to analyze the viability of castor seeds that were either unprepared or pretreated with chloroform, a gaseous contrast substance. However, they did not observe any significant differences in the estimates. The problem with using pretreatment is that seeds subjected to contrasts, in order to obtain a better resolution in X-ray images, generally lose their germination capacity, making *a posteriori* analysis with the same seed unviable (Gordon, Gosling, & Wang, 1991).

As noted, Bayesian point solutions and those obtained by the method of relative frequency showed very close values. However, incorporating uncertainty about these estimates is difficult in the frequentist context. The Bayesian method, on the other hand, offers flexibility in the inferential process and allows the estimation of parameters with associated levels of credibility. In addition, from a Bayesian perspective, it is possible to take advantage of previous knowledge to obtain more reliable combined estimates through meta-analyses, as well as to associate the researcher's experience by assigning informative priors, incorporating information from experiments prior to the analysis (Perez-Elizalde, Jarquin, & Crossa, 2011; Bodnar, Muhumuza, & Possolo, 2020).

Mixture distribution models have stood out in the modeling of phenomena whose population is composed of different subpopulations, similar to the problem addressed here. The Bayesian method with MCMC sampling offers many advantages in relation to other estimation procedures, such as a greater guarantee of convergence and the possibility of incorporating uncertainty through the estimated *a posteriori* distributions (Diebolt & Robert, 1994; Bilancia & Pollice, 1999; Stephens, 2000). In addition, when *a posteriori* distributions for unknown parameters are available, Bayesian methods provide valid inference without counting the assumption of asymptotic normality. This represents an advantage since the theory of the asymptotic maximum likelihood estimate (MLE) can be applied only to a large number of observations (McNeish, 2016).

Conclusion

The proposed classifications based on radiographic image evaluations allowed the association of seed morphological characteristics with viability, as well as the analysis of germination that was performed later. In this sense, associations of morphological attributes with non-germination were more conclusive. The Bayesian methodology used was efficient, thus allowing important information related to seed quality to be accessed through posterior probabilities and other amounts of interest with levels of uncertainty incorporated. In addition, the mixture model made it possible to properly associate seed probabilities in originating normal (or abnormal) seedlings with morphological characteristics.

Acknowledgements

The authors would like to thank the Federal University of Grande Dourados for their financial support and the Department of Agriculture of Federal University of Lavras for conducting the experiment.

References

- Achcar, J. A., & Ruffino Netto, A. (2003). Estudo da prevalência da tuberculose: uso de métodos bayesianos. *Revista Brasileira de Epidemiologia*, 6(4), 380-387. DOI: <https://doi.org/10.1590/S1415-790X2003000400012>
- Batista, T. B., Cardoso, E. D., Binotti, F. F. S., Costa, E., & Sá, M. E. (2016a). Priming and stress under high humidity and temperature on the physiological quality of *Brachiaria brizantha* cv. MG-5 seeds. *Acta Scientiarum. Agronomy*, 38(1), 123-127. DOI: <https://doi.org/10.4025/actasciagron.v38i1.26021>
- Batista, T. B., Binotti, F. F. S., Cardoso, E. D., Costa, E., & Nascimento, D. M. (2016b). Appropriate hydration period and chemical agent improve priming in brachiaria seeds. *Pesquisa Agropecuária Tropical*, 46(3), 350-356. DOI: <https://doi.org/10.1590/1983-40632016v4638422>
- Bilancia, M., & Pollice, A. (1999). Bayesian estimation of finite mixtures of Gaussian mixtures. *Metron - International Journal of Statistics*, 57(1-2), 181-208.
- Bino, R. J., Aartse, J. W., & Van Der Burg, W. J. (1993). Non-destructive X-ray analysis of *Arabidopsis* embryo mutants. *Seed Science Research*, 3(3), 167-170. DOI: <https://doi.org/10.1017/S0960258500001744>
- Bodnar, O., Muhumuza, R. N., & Possolo, A. (2020). Bayesian inference for heterogeneity in meta-analysis. *Metrologia*, 57(6). DOI: <https://doi.org/10.1088/1681-7575/abb064>
- Brasil. (2008). Ministério da Agricultura, Pecuária e do Abastecimento. Instrução Normativa nº 30, de 21 de maio de 2008. *Diário Oficial da União*, 23 maio 2008, Seção 1, p. 45.
- Brasil. (2009). Ministério da Agricultura, Pecuária e Abastecimento. *Regras para análise de sementes*. Brasília, DF: MAPA.
- Cicero, S. M, Van Der Heijden, G. W. A. M, Van Der Burg, W. J, & Bino, R. J. (1998). Evaluation of mechanical damage in seeds of maize (*Zea mays* L.) by X-ray and digital imaging. *Seed Science and Technology*, 26(3), 603-612.
- Cardoso, E. D., Sá, M. E., Haga, K. I., Binotti, F. F. S., Nogueira, D. C., & Valério Filho, W. V. (2014). Desempenho fisiológico e superação de dormência em sementes de *Brachiaria brizantha* submetidas a tratamento químico e envelhecimento artificial. *Semina: Ciências Agrárias*, 35(1), 21-38. DOI: <https://doi.org/10.5433/1679-0359.2014v35n1p21>
- Copeland, L. O., & McDonald, M. B. (2001). *Principles of seed science and technology* (4th ed.). New York, NY: Springer US.
- Diebolt, J., & Robert, C. P. (1994). Estimation of finite mixture distributions through Bayesian sampling. *Journal of the Royal Statistical Society*, 56(2), 363-375. DOI: <https://doi.org/10.1111/j.2517-6161.1994.tb01985.x>
- Gaspar-Oliveira, C. M., Martins, C. C., Nakagawa, J., & Cavariani, C. (2008). Duração do teste de germinação de *Brachiaria brizantha* cv. Marandu (Hochst. ex A. Rich.) Stapf. *Revista Brasileira de Sementes*, 30(3), 30-38. DOI: <https://doi.org/10.1590/S0101-31222008000300005>
- Gomes Junior, F. G. (2010). Aplicação da análise de imagens para avaliação da morfologia interna de sementes. *Informativo ABRATES*, 20(3), 33-51.
- Gomes, K. B. P., Martins, R. D. C. C., Martins, I. S., & Gomes Junior, F. G. (2014). Avaliação da morfologia interna de sementes de *Terminalia argentea* (Combretaceae) pelo teste de raios X. *Revista Ciência Agrônômica*, 45(4), 752-759. DOI: <https://doi.org/10.1590/S1806-66902014000400013>
- Gordon, A., Gosling, P., & Wang, B. (1991). *Tree and shrub seed handbook* (4th ed.). Zürich, SW: ISTA.
- Heidelberger, P., & Welch, P. (1983). Simulation run length control in the presence of an initial transient. *Operations Research*, 31(6), 1109-1144. DOI: <https://doi.org/10.1287/opre.31.6.1109>
- Jeromini, T. S., Martins, C. C., Pereira, F. E. C. B., & Gomes Junior, F. G. (2019). The use of X-ray to evaluate *Brachiaria brizantha* seeds quality during seed processing. *Revista Ciência Agrônômica*, 50(3), 439-446. DOI: <https://doi.org/10.5935/1806-6690.20190052>
- Junqueira, V. S., Peixoto, L. A., Laviola, B. G., Bhering, L. L., Mendonça, S., Costa, T. S. A., & Antoniassi, R. (2016). Bayesian multi-trait analysis reveals a useful tool to increase oil concentration and to decrease toxicity in *Jatropha curcas* L. *PLoS ONE*, 11(8), 1-2. DOI: <https://doi.org/10.1371/journal.pone.0161046>
- Lima, L. K., Santos, C. C., Dutra, A. S., & Barros, G. L. (2013). Utilização de técnicas na avaliação de sementes por imagem. *Agropecuária Científica no Semi-Árido*, 9(3), 01-06. DOI: <https://doi.org/10.30969/acsa.v9i3.291>

- Marin, J. M., Mengersen, K., & Robert, C. P. (2005). Bayesian modelling and inference on mixtures of distributions. *Handbook of Statistics*, 25, 459-507. DOI: [https://doi.org/10.1016/S0169-7161\(05\)25016-2](https://doi.org/10.1016/S0169-7161(05)25016-2)
- McNeish, D. (2016). On using Bayesian methods to address small sample problems. *Structural Equation Modeling: A Multidisciplinary Journal*, 23(5), 750-773. DOI: <https://doi.org/10.1080/10705511.2016.1186549>
- Medeiros, A. D., Pinheiro, D. T., Xavier, W. A., Silva, L. J., & Dias, D. C. F. S. (2020). Quality classification of *Jatropha curcas* seeds using radiographic images and machine learning. *Industrial Crops and Products*, 146, 1-7. DOI: <https://doi.org/10.1016/j.indcrop.2020.112162>
- Menezes, N. L., Cícero, S. M., & Villela, F. A. (2005). Identificação de fissuras em sementes de arroz, após a secagem artificial, por meio de raios X. *Ciência Rural*, 35(5), 1194-1196. DOI: <https://doi.org/10.1590/S0103-84782005000500033>
- Mondo, V. H. V., & Cicero, S. M. (2005). Análise de imagens na avaliação da qualidade de sementes de milho localizadas em diferentes posições na espiga. *Revista Brasileira de Sementes*, 27(1), 9-18. DOI: <https://doi.org/10.1590/S0101-31222005000100002>
- Novembre, A. D. D. L. C., Chamma, H. M. C. P., & Gomes, R. B. R. (2006). Viabilidade das sementes de braquiária pelo teste de tetrazólio. *Revista Brasileira de Sementes*, 28(2), 147-151. DOI: <https://doi.org/10.1590/S0101-31222006000200020>
- Pereira, C. E., Oliveira, J. A., Rosa, M. C. M., & Kikuti, A. L. P. (2011). Armazenamento de sementes de braquiária peletizadas e tratadas com fungicida e inseticida. *Ciência Rural*, 41(12), 2060-2065. DOI: <https://doi.org/10.1590/S0103-84782011001200004>
- Perez-Elizalde, S., Jarquin, D., & Crossa, J. (2011). A general Bayesian estimation method of linear-bilinear models applied to plant breeding trials with genotype×environment interaction. *Journal of Agricultural, Biological, and Environmental Statistics*, 17(1), 15-37. DOI: <https://doi.org/10.1007/s13253-011-0063-9>
- R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, AT: R Foundation for Statistical Computing.
- Raftery, A. E., & Lewis, S. (1992). How many iterations in the Gibbs sampler? In J. M. Bernardo, J. O. Berger, A. P. Dawid, & A. F. M. Smith (Eds.), *Bayesian Statistics* (p. 763-773). Oxford, UK: Oxford University.
- Rocha, R. C., Sáfiadi, T., & Carvalho, M. L. M. (2013). Análise bayesiana do potencial de viabilidade de sementes de mamona por meio de análise radiográfica. *Revista Brasileira de Estatística*, 74(238), 101-127.
- Rufo, M. J., Martín, J., & Pérez, C. J. (2006). Bayesian analysis of finite mixture models of distributions from exponential families. *Computational Statistics*, 21(3-4), 621-637. DOI: <https://doi.org/10.1007/s00180-006-0018-8>
- Ruggiero, M. A. G., & Lopes, V. L. R. (1997). *Cálculo numérico: aspectos teóricos e computacionais*. São Paulo, SP: Makron Books.
- Schlattmann, P. (2009). *Medical applications of finite mixture models*. Berlin, GE: Springer.
- Silva, G. Z., Martins, C. C., Cruz, J. O., Jeromini, T. S., & Bruno, R. L. A. (2017). Evaluation the physiological quality of *Brachiaria brizantha* cv. BRS 'Piatã' seeds. *Bioscience Journal*, 33(3), 572-580. DOI: <https://doi.org/10.14393/BJ-v33n3-36519>
- Simak, M., & Gustafsson, A. (1953). X-ray photography and sensitivity in forest tree species. *Hereditas*, 39(3-4), 458-468. DOI: <https://doi.org/10.1111/j.1601-5223.1953.tb03430.x>
- Simpson, D., Rue, H., Riebler, A., Martins, T. G., & Sorbye, S. H. (2017). Penalising model component complexity: a principled, practical approach to constructing priors. *Statistical Science*, 32(1), 1-28. DOI: <https://doi.org/10.1214/16-ST576>
- Smith, B. J. (2007). boa: An R package for MCMC output convergence assessment and posterior inference. *Journal of Statistical Software*, 21(11), 1-37. DOI: <https://doi.org/10.18637/jss.v021.i11>
- Stephens, M. (2000). Bayesian analysis of mixture models with an unknown number of components-an alternative to reversible jump methods. *The Annals of Statistics*, 28(1), 40-74. DOI: <https://doi.org/10.1214/aos/1016120364>
- Tanner, M. A., & Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, 82(398), 528-540. DOI: <https://doi.org/10.1080/01621459.1987.10478458>

Warnes, G. R., Bolker, B., & Thomas, L. (2018). *gtools: Various R Programming Tools*. R package version 3.8.1. Retrieved on Aug. 20, 2020 from <https://CRAN.R-project.org/package=gtools>