Fantazzini, Rosa, Carvalho, Liska, Carvalho, Coelho, Cirillo and Ribeiro (2020).
*Seed Science and Technology*, **48**, 2, 179-188.
https://doi.org/10.15258/sst.2020.48.2.05

# Correlation between historical data of the germination test and of the tetrazolium test in coffee seeds by GAMLSS

**Tatiana Botelho Fantazzini[1], Sttela Dellyzete Veiga Franco da Rosa[2*], Gladyston Rodrigues Carvalho[3], Gilberto Rodrigues Liska[4], Maria Laene Moreira de Carvalho[1], Stefânia Vilas Boas Coelho[1], Marcelo Ângelo Cirillo[5] and Fernando Augusto Sales Ribeiro[1]**

[1] Departamento de Agricultura, Universidade Federal de Lavras, UFLA, Caixa Postal 3037, 37200-000, Lavras, MG, Brazil (E-mail: tatiana_botelho@hotmail.com; mlaene@gmail.com; stefaniavbc@gmail.com)
[2] Empresa Brasileira de Pesquisa Agropecuária, Embrapa Café, Brasília – DF, Brazil
[3] Empresa de Pesquisa Agropecuária de Minas Gerais, Epamig Sul, Lavras – MG, Brazil (E-mail: carvalho@epamig.ufla.br)
[4] Universidade Federal do Pampa, (UNIPAMPA) – RS, Brazil (E-mail: gilbertoliska@hotmail.com)
[5] Departamento de Estatística, Universidade Federal de Lavras, Caixa Postal 3037, 37200-000, UFLA, Lavras, MG, Brazil (E-mail: macufla@gmail.com)
* Author for correspondence (E-mail: sttela.rosa@embrapa.br)

*(Submitted August 2018; Accepted January 2019; Published online April 2020)*

## Abstract

In the Brazilian National System of Seeds and Seedlings, coffee seeds are evaluated by the germination or tetrazolium test. However, differences have been observed between the results of these tests in various studies, especially when the seeds have a lower level of quality. Given this situation, the aim of this study was to evaluate the correlation between historical data of results of the germination test and of the tetrazolium test in samples of coffee seeds using Generalized Additive Models for Location, Scale and Shape (GAMLSS). Historical data of results of the germination test and of the tetrazolium test of coffee seeds originating from different cultivars and different crop seasons were used. The zero-or-one inflated beta GAMLSS is suitable for fitting data from the germination test and from the tetrazolium test. The estimate of viability by the tetrazolium test varies according to the germination percentage class. There are greater GAMLSS correlations between the percentages of normal seedlings and of viability in the tetrazolium test for germination values above 70%, and low correlations below this value, showing that evaluation of coffee seeds based only on the tetrazolium test may not correspond to actual physiological performance.

**Keywords**: germination potential, quality evaluation, regression analysis, viability

## Introduction

Brazilian coffee is an agricultural product of worldwide socioeconomic importance and coffee-growing is a large source of revenue for the Brazilian economy. Brazil is the largest producer and exporter of coffee (Abic, 2017; Conab, 2018).

The coffee crop is set up by means of seedlings originating from seeds. One of the critical points in coffee production is the use of quality seeds. Many steps and a great deal of care must be taken during seed production, which should conclude with evaluation of final seed quality so as to ensure high viability and vigour for formation of seedlings with the quality necessary for establishing vigorous and high-yielding crop fields.

To evaluate the physiological quality of coffee seeds, the germination test is specified (Brasil, 2009). However, due to problems such as the long time required for the test, together with the rapid loss of viability of coffee seeds, the results may not reflect the actual physiological state of the seeds (Dias and Silva, 1986). A feasible alternative is the tetrazolium test, which has been used instead of the germination test for coffee seeds because of its speed in determination of viability (Delouche *et al.*, 1976; França Neto *et al.*, 1988; Costa and Marcos Filho, 1994). The test is also easy to perform and evaluate (Sera and Miglioranza, 2003).

The efficiency of the tetrazolium test for evaluating seed viability depends on the development of methods appropriate for each species, determining adequate conditions for hydration, seed preparation, concentration of the tetrazolium solution, and the time and temperature of conditioning. These standardisation efforts are important because they directly influence the intensity and uniformity of the seed staining process and hence, the evaluation and interpretation of the results (Gaspar-Oliveira *et al.*, 2009; Souza *et al.*, 2010).

Determination of these parameters generally occurs based on studies that compare the results obtained in the tetrazolium test with the different tests for evaluation of seed quality, such as the germination test and the field seedling emergence test (Pasha and Das, 1982; Barros and Marcos Filho, 1990; Gaspar-Oliveira *et al.*, 2009).

The tetrazolium test can be used for analysis of coffee seed viability in accordance with Normative Instruction IN – no. 35 of 29 November 2012 of the Brazilian Ministry of Agriculture (Ministério da Agricultura, Pecuária e Abastecimento – MAPA) (Brasil, 2012) and as established by the Rules for Seed Analysis (Regras para Análise de Sementes - RAS) (Brasil, 2009) as an alternative for verifying the germination potential of seed lots. Despite frequent use of the tetrazolium test for coffee seeds, there are still discrepancies with the results of the germination test; authors have observed a tendency of overestimating, or in some cases underestimating, the germination potential of seed lots (Clemente *et al.*, 2011; Coelho *et al.*, 2015; Figueiredo *et al.*, 2017). Thus, more studies should be performed to better understand the correlation between the values of germination potential estimated by the tetrazolium test and the percentages of normal seedlings obtained through the germination test so as to clarify possible variations in these results.

Statistical modeling of historical values like those of coffee seed germination and viability constitutes quite a complex task since, considering a historical database, many observations corresponding to 0 and 100% germination are found. This results in statistical distributions with a high degree of asymmetry, and the usual methods of inference, such as simple linear regression or analyses of variance, become unreliable techniques for comparison of means and modeling of data. In addition, multiple associated factors, such as different crop seasons, cultivars, non-linear effects and heterogeneous variances for different levels (values) of the factors must be considered.

Such data do not conform to the normal distribution; therefore, more complex methods that allow these effects to be accommodated are necessary so that predictions related to the germination from the tests are reliable and accurate. An alternative method involves fitting Generalized Additive Models for Location, Scale and Shape (GAMLSS) that are suitable for fitting data with asymmetric and flexible characteristics. These models support high rates of null observations or extra heterogeneity among sampling units, as for example, the zero-or-one inflated beta distribution.

Therefore, the aim of this study was to propose modeling that uses GAMLSS models adapted to the zero-or-one inflated beta distribution so that the correlation between the historic data of the results of the germination test and of the tetrazolium test in coffee seeds can be explained in accordance with different classes of germination percentage.

## Materials and methods

The study was conducted in the Central Seed Laboratory of the Department of Agriculture of the Universidade Federal de Lavras using historical data of the results of the germination test and of the tetrazolium test of coffee seeds originating from different agricultural crop seasons and cultivars. Results from 548 coffee seed samples were used, part of them corresponding to samples received in the Seed Analysis Laboratory for evaluation and for issuing reports for the certification system, and part originating from research projects performed within the sphere of the Graduate Studies Program in Plant Science/Agronomy of the Universidade Federal de Lavras. All the seed samples were prepared and analysed using the same method in the germination and tetrazolium tests, as described.

The parchment (the part of the fruit that remains on the seeds) was manually removed from the seeds before the tetrazolium and germination tests were performed. The seeds were soaked in water for 48 hours at 30°C, after which the embryos were extracted manually, according to the Rules for Seed Analysis (Brasil, 2009).

*Tetrazolium test*

The tetrazolium test was performed with four replications of 50 seeds for each treatment; the seeds were soaked in water for 48 hours to allow extraction of the embryos. The embryos were placed in a polyvinylpyrrolidone (PVP) antioxidant solution and, after extraction, the embryos were washed in running water and immersed in 0.5% tetrazolium solution in dark-coloured bottles, which were kept at 30°C for three hours (Brasil, 2009; Clemente *et al*., 2011). Embryo viability was analysed with the aid of a stereo magnifier with 10 magnification for better visualisation of the internal and external structures, and seeds were then classified as viable or inviable according to the location and extent of the damage (Brasil, 2009).

*Germination test*

The germination test was performed with four replications of 50 seeds for each treatment, which were sown in rolls of "germitest" paper toweling moistened with distilled water in the amount of 2.5 times the weight of the dry paper. The seeds were sown and placed in a germinator regulated to 30°C and in the presence of light (Brasil, 2009). The percentage of normal seedlings at 30 days after sowing was determined; normal seedlings were considered those that had a primary root and at least two lateral roots.

*Description of the generalized additive model for location, scale and shape (GAMLSS) applied in evaluation of the viability of coffee seeds by the germination and tetrazolium tests*

The Generalized Additive Models for Location, Scale and Shape (GAMLSS) originally proposed by Rigby and Stasinopoulos (2005) were used to analyse the correlation between the germination and tetrazolium results.

These models belong to a more flexible class of models that consider mean value, dispersion and symmetry of the data in the modeling process when fitting models that envisage the nature of the response. The responses in this study referred to proportions limited within the interval [0,1], and a large number of seed samples exhibited null observations, i.e., zero percent germination, whereas in the tetrazolium test, the corresponding values of viability were greater than zero. Therefore, the use of more complex modeling became necessary, characterised by the zero-or-one inflated beta distribution, adapted to the GAMLSS models, a description of which follows.

*GAMLSS model for zero-or-one inflated beta distribution*

The description of this model begins by considering the mixture between the beta and Bernoulli distributions. According to Ospina and Ferrari (2010), the probability density function of the zero-or-one inflated beta distribution (INFBE) is given by

$$f\left(y;\alpha,\gamma,\mu,\sigma\right)=\begin{cases}\alpha\left(1-\gamma\right), & y=0\\ \alpha\gamma, & y=1\\ \left(1-\alpha\right)f\left(y;\mu,\sigma\right), & y\in\left(0,1\right)\end{cases} \tag{1}$$

with $0<\alpha, \gamma, \mu<1$ and $\sigma>0$, in which $f(y;\mu,\sigma)$ is the beta density function, given by

$$f\left(y;\mu,\sigma\right)=\frac{\Gamma\left(\sigma\right)}{\Gamma\left(\mu\sigma\right)\Gamma\left(\left(1-\mu\right)\sigma\right)}y^{\mu\sigma-1}\left(1-y\right)^{(1-y)\sigma-1} \tag{2}$$

in which $\Gamma(\cdot)$ is the gamma function. In addition, the GAMLSS model considers a vector of observations of response variable $\mathbf{y}^T=(y_1, y_2,..., y_{755})$, with $n=755$ samples evaluated, $g_k(\cdot)$, $k=1,...,4$, a monotonic link function relating $\theta_k$ to the coffee seed sample, composed of 548 samples, and the linear predictor of the additive model is given by

$$g_k\left(\theta_k\right) = \eta_k = \sum_{j=1}^{J_k} h_{jk}\left(x_{jk}\right) \tag{3}$$

in which $\theta_k$ and $\eta_k$ are vectors of size $n$, $\beta_k^T = (\beta_{1k}, \beta_{2k}, ..., \beta_{J_k k})$ is a vector of size, $J_k$, $\mathbf{x}_k$ is a known design matrix of size, $n \times J_k$, and $h_{jk}$ $(x_{jk})$ is an unknown function of the explanatory variable $x_{jk}$. In the present study, $h_{jk}$ $(x_{jk})$ is a penalised beta spline, which is characterised by polynomials defined by parts according to a beta spline base function in the explanatory variable. The coefficients of the base functions are penalised to ensure sufficient smoothness (Eilers and Marx, 1996). The model in (3) is denominated semiparametric GAMLSS. Thus, it will be assumed that the germination rate in the different methods has zero-or-one inflated beta distribution, i.e., $\mathbf{y} \sim BEINF$ $(\mu, \sigma, \nu, \tau)$, with a linear predictor for the parameters $\mu$ and $\sigma$, denominated location and scale parameters, and $\nu$ and $\tau$ denominated shape parameters, given by

$$\begin{cases} g_1\left(\mu\right) = \eta_1 = \operatorname{logit}\left(\mu\right) = \sum_{j=1}^{J_1} h_{j1}\left(x_{j1}\right) \\ g_2\left(\sigma\right) = \eta_2 = \operatorname{logit}\left(\sigma\right) = \sum_{j=1}^{J_1} h_{j2}\left(x_{j1}\right) \\ g_3\left(\nu\right) = \eta_3 = \log\left(\nu\right) = \sum_{j=1}^{J_1} h_{j3}\left(x_{j1}\right) \\ g_4\left(\tau\right) = \eta_4 = \log\left(\tau\right) = \sum_{j=1}^{J_1} h_{j4}\left(x_{j1}\right) \end{cases} \tag{4}$$

The parameters of the model were simultaneously estimated from the RS back-fitting algorithm (Rigby and Stasinopoulos, 2007). This algorithm has the advantage of not needing initial estimates of the values of each one of the parameters associated with the predictor variables. Only initial estimates of the values of the original parameters of the distribution of the response variable are necessary.

The worm plot was used as a diagnostic tool of the fit of the proposed model. According to Buuren and Fredriks (2001), the worm plot consists of ordered pairs formed by the quantile of standard normal distribution and the difference between the empirical quantile of the residual of the proposed model and standard normal. The way the ordered pairs are arranged indicates the suitability of the model.

The fit of the model and the goodness of fit diagrams were obtained through the R program (R Core Team, 2017) with resources from the *gamlss* package (Rigby and Stasinopoulos, 2017).

## Results

Figure 1 shows the results of fitting inflated beta GAMLSS models to the germination data (proportion of normal seedlings) and the tetrazolium data (proportion of viable embryos).

Good fit did not occur for all the observations. Thus, fitting a single model to all the data is not sufficient to explain the proportion of germinated or viable seeds. Consequently, refitting the model to the two sets of stratified test results was carried out. Figures 2 and 3 show the inflated beta GAMLSS model fitted to the germination and tetrazolium data within each class of predicted proportions: [0,10], [11,20], [21,30], [31,40], [41,50], [51,60], [61,70], [71,80], [81,90] and [91,100]. For visualisation in the graphs and fit of the model, the germination data were divided by 100 to respect the condition of the model that germination data are in the closed interval [0,1].

From the models fitted in classes [0,10], [11,20], [21,30], [31,40], [41,50] and [51,60] (figure 2), in general, in the tetrazolium test, the percentage of viable embryos is higher than the percentage of normal seedlings in the germination test. Beginning with class [61,70], mean values of the normal seedling variable and viable embryo variable were ever nearer, especially for the values greater than 70% (figure 3), the minimum standard for sale of coffee seeds.

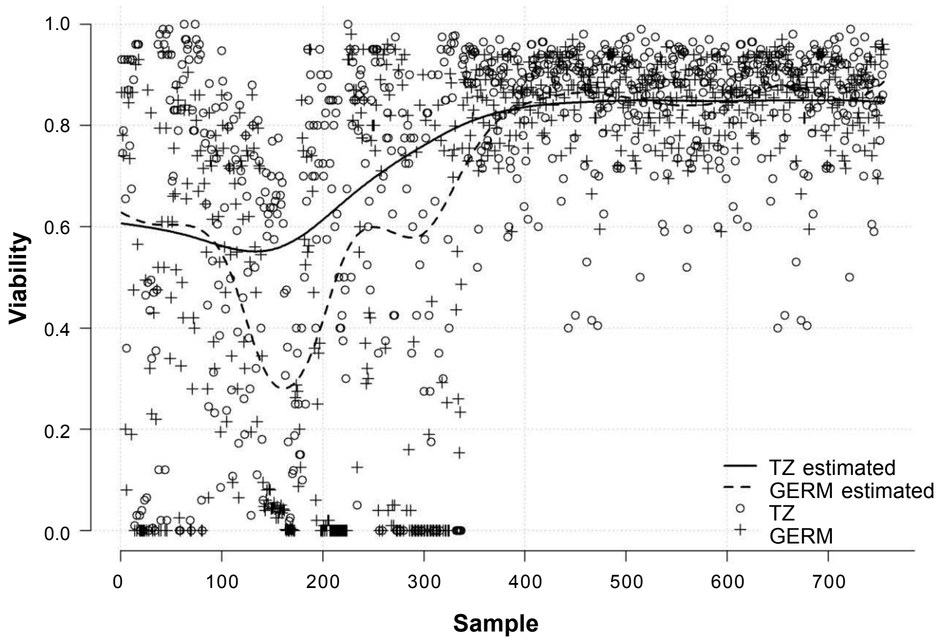


Figure 1. Inflated beta GAMLSS models fitted to the data of evaluation for the variables of normal seedlings and viable embryos by the tetrazolium test of *Coffea arabica* seeds.

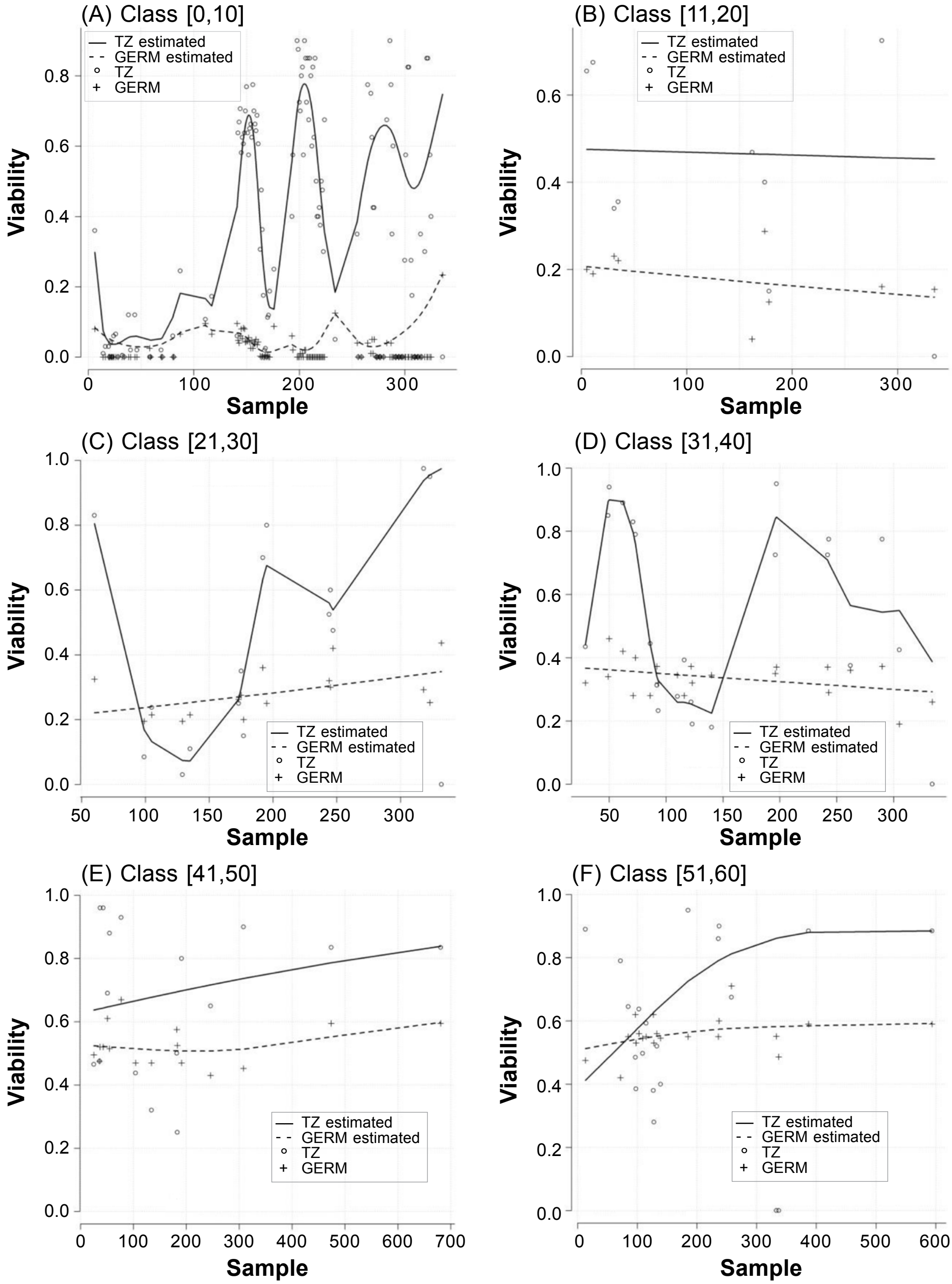


Figure 2. Inflated beta GAMLSS model fitted to the proportion of germinated seeds (normal seedlings) and proportion of viable seeds (embryos stained by the tetrazolium test) of *Coffea arabica.* Data were stratified to classes of predicted proportions as indicated. Data were transformed to scale 0-1.

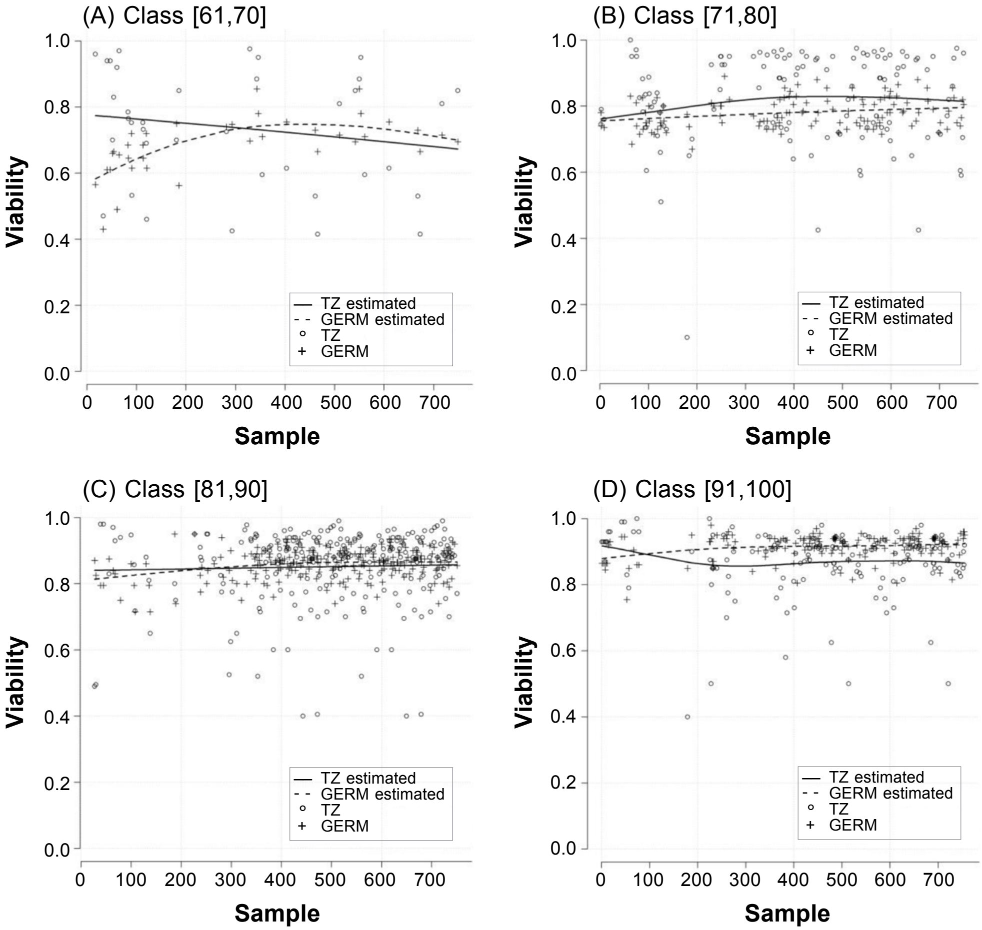


Figure 3. Inflated beta GAMLSS model fitted to the proportion of germinated seeds (normal seedlings) and proportion of viable seeds (embryos stained by the tetrazolium test) of *Coffea arabica*. Data were stratified to classes of predicted proportions as indicated. Data were transformed to scale 0-1.

## Discussion

Various studies have confirmed the discrepancy between the results of the tetrazolium test and of the germination test in coffee seeds, especially in lower quality seeds (Dussert and Engelmann 2006; Coelho *et al*., 2015; Figueiredo *et al*., 2017).

From these results, it can be confirmed that the potential of the tetrazolium test in estimating the quality of seeds varies according to the class of germination values. Notably, in seeds classified with germination lower than 60%, the results of the tetrazolium test overestimate the germination values, i.e., it is not a good test for evaluating the quality of these seeds. However, for seeds with germination above 60%, the results of the tetrazolium test were near the values of the germination test, indicating that for these seeds, the tetrazolium test can be used to estimate seed quality.

Analysis of correlation from GAMLSS shows that there is a good correlation between the data of germination and of viability for values above the minimum standard for sale of coffee seeds (Brasil, 2009), i.e., above 70%. Nevertheless, it also shows that evaluation of seeds based only on the tetrazolium test may recommend approval of seed lots for sale of coffee seeds whose actual value of germination is well below the minimum.

Aiming to validate the analyses of correlation by the GAMLSS model, residual analysis was performed. According to the correlation of the residues in both tests, it can be confirmed that the zero-or-one inflated beta GAMLSS regression model did not show serious evidence of lack of fit, since most is contained in the respective confidence intervals.

## Acknowledgements

Our thanks to Embrapa Café, to the Coffee Research Consortium, to the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), to the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), to the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), and to the Universidade Federal de Lavras (UFLA).

## References

Associação Brasileira da Indústria de Café (2017). *Produção Mundial de Café - Principais Países Produtores*, Safra. [*World Coffee Production – Major Producing Countries*]. Available at: <http://www.abic.com.br/publique/cgi/cgilua.exe/sys/start.htm?sid=48#2810>. Accessed on: Oct. 25th, 2017.

Barros, A.S.R. and Marcos Filho, J. (1990). Testes para avaliação rápida da viabilidade de sementes de soja. [Tests for rapid assessment of the viability of soybean seeds]. *Pesquisa Agropecuária Brasileira*, **25**, 1447-1459.

Brasil (2009). *Regras para Análise de Sementes.* [*Rules for Seed Analysis*], Ministério da Agricultura, Pecuária e Abastecimento, Secretaria de Defesa Agropecuária, Brasília.

Brasil (2012). Ministério da Agricultura, Pecuária e Abastecimento. Instrução Normativa nº 35, de 29 de novembro de 2012. [Normative Instruction nº 35, November 29th, 2012]. Available at: <http://www.lex.com.br/legis_24015030_ Instrucao_Normativa_N_35_de_29_de_Novembro_de_2012.aspx>. Accessed on: Aug. 28th, 2017.

Buuren, S.V. and Fredriks, M. (2001). Worm plot: a simple diagnostic device for modelling growth reference curves. *Statistics in Medicine*, **20**, 1259-1277.

Clemente, A.C.S., Carvalho, M.L.M., Guimarães, R.M. and Zeviani, W.M. (2011). Preparo das sementes de café para a avaliação da viabilidade pelo teste de tetrazólio. [Preparation of coffee seeds to assess viability using the tetrazolium test]. *Revista Brasileira de Sementes*, **33**, 38-44. <http://doi.org/10.1590/S0101-31222011000100004>

Coelho, S.V.B., Figueiredo, M.A., Clemente, A.C.S., Coelho, L.F.S. and Rosa, S.V.F. (2015). Alterações fisiológicas e bioquímicas em sementes de café secas em sílica gel e soluções salinas saturadas. [Physiological and biochemical changes in coffee seeds dried in silica gel and saturated saline solutions]. *Pesquisa Agropecuária Brasileira*, **50**, 483-491. <http://doi.org/10.1590/S0100-204X2015000600007>

Conab (2018). *Acompanhamento da Safra Brasileira de café, Safra 2018: Primeiro Levantamento* [*Follow-up of the Brazilian Coffee Crop, 2018 Harvest: First Survey*], Companhia Nacional de Abastecimento, Brasília. Available at: <https://conab.gov.br/info-agro/safras/cafe>.

Costa, N.P. and Marcos Filho, J. (1994). O emprego do teste de tetrazólio na avaliação da qualidade da semente de soja. [The use of the tetrazolium test in the evaluation of soybean quality]. *Informativo ABRATES*, **4**, 53-62.

Delouche, J.C., Still, T.W., Raspet, M. and Lienhard, M. (1976). O teste de tetrazólio para viabilidade da semente (James C. Delouche), [The tetrazolium test of seed viability] AGIPLAN, Brasília.

Dias, M.C.L.L. and Silva, W.R. (1986). Determinação da viabilidade de sementes de café através do teste de tetrazólio. [Determination of the viability of coffee seeds through the tetrazolium test]. *Pesquisa Agropecuária Brasileira*, **21**, 1139-1145.

Dussert, S. and Engelmann, F. (2006). New determinants for tolerance of coffee (*Coffea arabica* L.) seeds to liquid nitrogen exposure. *CryoLetters*, **27**, 169-178.

Eilers, P.H.C. and Marx, B.D. (1996). Flexible smoothing with B-splines and penalties. *Statistical Science*, **11**, 89-121.

Figueiredo, M.A., Coelho, S.V.B., Rosa, S.D.V.F., Vilela, A.L. and Silva, L.C. (2017). Exploratory studies for cryopreservation of *Coffea arabica* L. seeds *Journal of Seed Science*, **39**, 150-158. <http://doi.org/10.1590/2317-1545v39n2171123>

França Neto, J.B., Pereira, L.A., Costa, N.P., Krzyzanowski, F.C. and Henning, A.A. (1988). *Metodologia do Teste de Tetrazólio em Sementes de Soja*, [*Methodology of the Tetrazolium Test in Soybean Seeds*], EMBRAPA-CNPSo, Londrina.

Gaspar-Oliveira, C.M., Martins, C.C. and Nakagawa, J. (2009). Concentração da solução de tetrazólio e período de coloração do teste para sementes de mamoneira. [Tetrazolium solution concentration and test staining period for castor bean seeds]. *Revista Brasileira de Sementes*, **31**, 38-47. <http://doi.org/10.1590/S0101-31222009000300004>

Ospina, R. and Ferrari, S.L.P. (2010). Inflated beta distributions. *Statistical Papers*, **51**, 111-126.

Pasha, M.R. and DAS, R.K. (1982). Quick viability test of soybean seed by using tetrazolium chloride. *Seed Science and Technology*, **10**, 651-655.

R Core Team (2017). R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Rigby, R.A. and Stasinopoulos, D.M. (2005). Generalized additive models for location, scale and shape. *Journal of Applied Statistics*, **54**, 507-554. <http://doi.org/10.1111/j.1467-9876.2005.00510.x>

Rigby, R.A. and Stasinopoulos, D.M. (2007). Generalized additive models for location scale and shape (GAMLSS) in R. *Journal of Statistical Software*, **23**, 1-46.

Rigby, R.A. and Stasinopoulos, D.M. (2017). Distributions for modelling location, scale and shape: using (GAMLSS) in R. Available at: <https://www.gamlss.com/wp-content/uploads/2018/01/DistributionsForModellingLocationScaleandShape.pdf>. Accessed on: Nov. 21th, 2017.

Sera, G.H. and Miglioranza, É. (2003). Avaliação visual do potencial germinativo de sementes de café pelo formato e coloração do embrião. [Visual evaluation of germinative potentiality of coffee seeds by embryo format and coloration]. *Semina: Ciências Agrárias*, **24**, 307-310.

Souza, C.R., Ohlson, O.C. and Panobianco, M. (2010). Avaliação da viabilidade de sementes de aveia branca pelo teste de tetrazólio. [Evaluation of white-oat seed viability by the tetrazolium test]. *Revista Brasileira de Sementes*, **32**, 174-180. <http://doi.org/10.1590/S0101-31222010000400020>