Nonlinear models to describe the maize seed quality during the maturation stage: a Bayesian approach

Sebastião Gazola[1], Carlos Alberto Scapim[2], Ângela Maria Marcone de Araujo[1], Robson Marcelo Rossi[1], Antonio Teixeira do Amaral Júnior[3*], Marcelo Vivas[3]

[1]Department of Statistics, Universidade Estadual de Maringá, 87020-900, Maringá, PR, Brazil
[2]Department of Agronomy, Universidade Estadual de Maringá, 87020-900, Maringá, PR, Brazil
[3]Laboratory of Plant Breeding, Universidade Estadual do Norte Fluminense Darcy Ribeiro, 28013-602, Campos dos Goytacazes, RJ, Brazil

*Corresponding author: amaraljr@pq.cnpq.br

Abstract

The adjustment of linear and non-linear models to describe the longevity of seed was studied here. The Bayesian analysis is a robust statistical procedure with many possible applications. In this study, the Bayesian method was used to fit the seed germination data of two maize hybrids (OC705 and CD5501) as a function of the number of days after female flowering on two sowing dates (E1 and E2) to the following non-linear model: $y(t) = A - B\exp(-Ct)$. The accumulated dry biomass was also fit to the following nonlinear model: $y(t) = A/(1 + \exp((B - t)/C)$. Ten consecutive assessments corresponding to intervals of four days were performed with four replicates. The sampling started 23 days after female flowering and the last sample was collected 59 days after this period. The Bayesian method enabled the study of germination curves. The results suggested the first sowing date (E1) for both hybrids as the most viable for planting. The results indicated 57 days after the female flowering as the ideal time for harvesting, when the maximum germination percentage (96 %) occurred. The procedure used in the present study also allowed an accurate comparison through the credibility intervals of the difference of the fit parameters, the hybrid types and the sowing dates. Therefore, this method was successfully applied to the seed industry to study germination percentage.

Keywords: Zea mays L., seed germination, female flowering.
Abbreviations: CrIs_credibility intervals; Dif_parameter differences; Est_estimated; MCMC_Markov chain Monte Carlo; Obs_observed.

Introduction

The fitting of linear and nonlinear models to describe the longevity of seed lots has been the target of various research groups (Andreoli and Andrade, 2007; Marcondes et al., 2011). Many of these models are based on Probit, linear and nonlinear regressions (Sanchez-Coronado et al., 2007). In contrast, Bayesian inference is not yet common to describe germination percentages of maize seeds. However, Filho et al. (2008) tested parameter consistency based on credibility intervals of growth curves fit to data from common bean cultivars. The advantages of use of Bayesian framework in the analysis of genetic experiments have been presented and discussed by several authors (Mora et al., 2009; Cané-Retamales et al., 2011; Safner et al., 2011; Silva et al., 2013; Mora and Serra, 2014).

In maize, Balestre et al. (2012a) conducted a Bayesian survey of multiple corn characteristics, in which they highlighted the importance of pleiotropic effects in the study of the inheritance of quantitative characteristics. The Bayesian analysis was used to study genetic control of resistance to gray leaf spot in maize (Balestre et al., 2012b), and in the molecular analysis of genetic structure of a Brazilian popcorn germplasm (Saavedra et al., 2013).

The Bayesian method considers the levels of uncertainty in the estimated parameters and the credibility intervals are generally more accurate than the confidence intervals, obtained using frequentist inference (Sorensen and Gionola, 2002). This is a robust statistical procedure with many possible applications (Kass and Rafetry, 1995; Meuwissen et al., 2001; Silva et al., 2013). In Bayesian inference, *a priori* information about the parameters is used in association with the sample data by the likelihood function generating a joint *a posteriori* distribution (posteriori $\propto$ likelihood x *priori*) (Geweke, 1992; Kass and Rafetry, 1995; Sorensen and Gionola, 2002). Thus, the conditional marginal distributions of the parameters obtained from the integration of the joint *a posteriori* distribution provide the estimators of interest (Sorensen and Gionola, 2002).

Considering the importance of germination percentage of seeds, this paper aims to: determine the best harvest time to get maximum germination percentage and the maximum time to obtain the highest cumulative dry biomass of two hybrids in two seasons seeding by Bayesian inference.

Results and Discussion

Study of dry biomass

Fig.1 shows the observed (Obs) and estimated (Est) dry biomass accumulation data for the OC705 and CD5501 hybrids sown on both dates (E1 and E2). The OC705 hybrid had different dry biomass accumulations when sown at times E1 and E2, and the same result observed in CD5501 hybrid. Moreover, the OC705 and CD5501 hybrids had similar observed and fit curves at time E1, and they also had similar curves at time E2.

The credibility interval for the difference of the parameter estimates indicated that the dry biomass of the OC705 and CD5501 hybrid seeds at time E1 or time E2 were not significantly different. However, the OC705 hybrid had significantly different accumulated dry biomass when comparing the two sowing dates and the same result occurred with the CD5501 hybrid. The distribution of dry biomass accumulation data was similar for the different hybrids sown on the same date, but these data were different for the same hybrid sown on different dates (Table 1).

The highest cumulative dry biomass was at 48 and 50 days for both seeded hybrid at the time E1. This result differed from those found in the works of Araujo et al. (2006a) for sweet corn cultivar sweet crystal, in which the maximum dry matter occurred at approximately 41 days after flowering. However the results of these authors are similar to the results found in this work for the time E2 (Table 3), where the maximum accumulated dry biomass occurred after 39 and 40 days. The longer time to reach the maximum of dry biomass in this study was 50 days at the E1, differing from the work of Borba et al. (1994) who studied the dry matter in simple hybrid maize seed BR 201 female. They observed that the maximum dry mass occur until 65 days after flowering, when produced in winter. Borba et al. (1995) found that the dry matter accumulation was stabilized at 58 days after flowering. David et al. (2002) reported that the maximum dry matter was achieved 68 days after flowering for DFT-2 variety of popcorn. Our results showed that hybrid corn seeds OC705 and CD5501 in E1 and E2 periods, respectively, acheived higher dry mass than simple seed corn BR 201 female and popcorn.

Study of germination percentage

The seed germination data were fit to the following exponential model: $y(t) = A - B \exp(-Ct)$. The parameter estimates are shown in Table 2. Parameter A was the maximum germination, and C was the growth rate of the germination percentage over the sampling time. Fig 2 shows the curves of the observed and fit data of the OC705 and CD5501 maize hybrids sown at times E1 and E2.

The credibility intervals (CrIs) allowed a comparison of the estimated parameters in the same manner, when Filho et al. (2008) compared growth curves of bean plants. The zero value belonging to the CrI indicated non-significant differences for the maximum germination parameters estimated for both hybrids and sowing dates, i.e., the germination percentages of these two hybrids and at both sowing dates were similar. Similar responses were observed for the growth rate parameter of the germination percentage. Thus, the equations fit for the seeds of the OC705 and CD5501 hybrids on the 28 October sowing date were similar, and the same trend occurred for the 20 November sowing date. However, the growth rate of the OC705 hybrid was superior at the first sowing date (Table 2).

For the OC705 hybrid, the B parameters were significantly different (zero did not belong in the CrI) between 28 October and 20 November, indicating that the germination percentages of this hybrid on the two sowing dates were significantly different (Table 2).

The time for the maximum germination of seed (Table 3) was similar for all hybrids planted at two periods. The CD5501 sown on October 28 had the highest percentage (96.2%) in 57 days after the female flowering, followed by OC705 which showed 95% at 56 days. The CD5501 had the lowest percentage when sown on 20 November. The best planting date was October 28 and 57 days after female flowering was the best harvest time for both hybrids. Similar results were obtained by Borba et al. (1994) for the simple hybrid maize BR 201 (*Zea mays* L.), which concluded that the harvest should begin 55 after flowering, when the seeds were about 87% germination. For sweet corn seeds (Doce Cristal cultivar) Araujo et al. (2006a) found that the maximum germination occurs at 55 days after flowering when the germination reaches up to 97%.

Borba et al. (1995) studied the seed of simple hybrid maize BR 201 female and reported 98.8% of germination at 44 days after flowering in winter. David et al. (2002) obtained the maximum germination by 99% at 44 days after flowering for popcorn, DFT-2 variety. Araujo et al. (2006b) reported that for sweet corn seeds harvested at different times, the harvest should be held to 47 days after flowering reaching a maximum of 88% germination. They reported a shorter time to harvest compared to the findings of this study for OC705 and CD5501 hybrid, which were 28 Oct and 20 Nov, respectively.

In this work, it was possible to determine the appropriate time through Bayesian inference, at which the plants should be harvested, i.e., the moment when the maximum germination percentage is reached. This percentage is indicative of the maximum physiological quality, and it helps to avoid unnecessary permanence time in the field, which frequently compromises seed performance. The procedure used in the present study also allowed an accurate comparison through the credibility intervals of the difference of the fit parameters, the hybrid types and the sowing dates. Therefore, this method was successfully applied to the seed industry to study germination percentage.

Materials and Methods

Plant materials

The field experiment was conducted at the Research Station of the Central Cooperative of Agricultural Research (Cooperativa Central de PesquisaAgrícola - Coodetec) in Cascavel, Brazil (latitude 24°56'S, longitude 53°26'W and altitude 760 m) in dystrophic dusky red latosol.

The field experiment consisted of two sowing dates on 28th Oct (E1) and 20th Nov (E2) in 2006 with two maize hybrids, namely, OC705 and CD5501. When the ears started to appear, the plants uniform in appearance were selected and tagged for sampling later. The day, in which 75% of the tagged plants in the plots showed 2.5 cm long silks, was considered the first day of female flowering. A completely randomised block design with four replications was used. The corn hybrids were seeded in the plots and, in the subplots, 10 samples were consecutively taken in regular periods of four days (split-plot in time). The sampling was initiated 23 days after female flowering (DAFF) and the last sample was collected 59 DAFF.
Each plot was composed of four twelve-meter lines,

Table 1. Estimates of parameter differences (Dif) fit by equation (2) to dry biomass accumulation during the maturation of the OC705 and CD5501 hybrid seeds sown on 28 October (E1) and 20 November (E2).

| Hybrid Time	Parameter differences	Mean	Standard Deviation	Error Monte Carlo	CrI(2,5%)	CrI(97,5%)
OC705	Dif A	0.071	16.850	0.114	-33.040	33.010
CD5501	Dif B	0.401	1.162	0.008	-1.902	2.655
E1	Dif C	0.804	1.403	0.010	-1.936	3.573
OC705	Dif A	-0.132	16.790	0.169	-33.130	32.400
CD5501	Dif B	0.371	1.150	0.009	-1.902	2.586
E2	Dif C	0.772	1.419	0.012	-1.962	3.565
OC705	Dif A	0.241	17.98	0.137	-35.360	35.820
E1	Dif B	6.009	1.437	0.012	3.206	8.817
E2	Dif C	1.695	1.317	0.010	-0.868	4.282
CD5501	Dif A	0.210	18.36	0.134	-35.801	35.702
E1	Dif B	6.001	1.44	0.011	3.174	8.816
E2	Dif C	1.691	1.32	0.010	-0.874	4.267
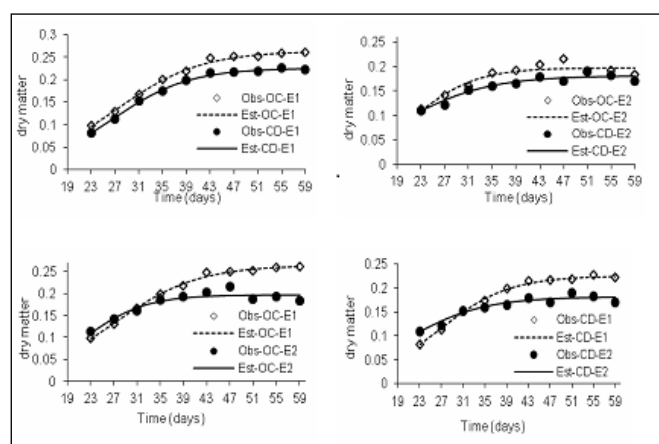
Fig 1. Observed (Obs) and estimated (Est) dry biomass accumulation curves for the OC705 and CD5501 hybrids sown on two different dates (E1, 28 October; and E2, 20 November).

where the first and the last lines, as well as the meter-end portion of the second and third ones were considered as borderlines. Thirty-four ears were sampled at each time. From each ear, 20 to 25 seeds were collected in the middle portion from a single row. After the ear was sampled for the first time, it was covered with a paper bag and tied with a rubber band, to prevent access to the ear through the opening where seeds were collected. This procedure was repeated after each ear sampling. Five seeds at random from each ear were taken to evaluate the black layer development and the milk line position.

Measured characteristics

The remaining seeds were put together as one sample and taken to the laboratory for evaluation of seed moisture content and seed germination percentage, according to the Brazilian Official Seed Analysis Rules (Brasil, 1992), dry seed weight and seed vigour using the cold and the accelerated aging tests. From time to time, with intervals of four days, seeds from the borderlines and at the same age as the ones being investigated were collected and their milk lines compared to verify the stage of development of the seeds remaining in each ear. The environmental conditions under which the field tests were conducted are summarizes in Fig 1.

Statistical analysis

The seed germination percentage was described with the empirical asymptotic nonlinear model (Ratkowsky, 1983) as follows:

$$y_i = A - B.\exp(-Ct_i) + \varepsilon_i \quad i=1, \quad 2,..., \quad n \quad (1)$$

Where; y is the seed germination percentage, A is the maximum value that the response y can reach, C is the growth rate, B is a constant inherent to the phenomenon under study and ε represents the random error.

Assuming that the residuals from the model described in equation (1) are additive, independent, and normally distributed with zero mean and variance (σ_e^2), the following equation is obtained:

$$(y_i \mid \sigma_e^2) \sim N(A - B.\exp(-Ct_i), \sigma_e^2)$$

The following *a priori* distributions were considered for the parameters of the model: A ~ N(0;10⁻⁴), B ~ Gamma(10⁻¹;10⁻³), C ~ Gamma(10⁻¹;10⁻³) and τ ~ Gamma(10⁻³;10⁻³) such that $\sigma_e^2 = \tau_e^{-1}$ (OpenBUGS parameterization).

Table 2. Estimates of parameters and parameter differences (Dif) fit by equation (1) to the germination data of the OC705 and CD5501 hybrid seeds sown on 28 October (E1) and 20 November (E2).

| Hybrid Time	Parameters	Mean	Standard Deviation	Error Monte Carlo	CrI(2,5%)	CrI(97,5%)
OC705 E1	A	0.9705	0.004058	5.81E-05	0.9626	0.9785
	B	11.26	0.5179	0.006271	10.3	12.33
	C	0.1124	0.001999	2.08E-05	0.1085	0.1164
CD5501 E1	A	1.000	0.006477	8.05E-05	0.9951	1.021
	B	7.576	0.3715	0.004849	6.883	8.328
	C	0.09295	0.002192	2.06E-5	0.08865	0.09723
OC705	Dif A	-0.0569	20.46	0.1488	-40.17	39.84
CD5501	Dif B	3.653	2.138	0.01628	-0.5162	7.868
E1	Dif C	-0.2421	31.61	0.2361	-61.94	61.26
OC705 E2	A	1.000	0.005061	5.81E-05	0.9978	1.018
	B	4.722	0.1689	0.002117	4.401	5.065
	C	0.08081	0.001655	2.08E-05	0.07761	0.08409
CD5501 E2	A	1.000	0.006466	8.05E-05	0.9932	1.019
	B	4.31	0.138	0.001838	4.05	4.59
	C	0.07186	0.001543	2.06E-5	0.06886	0.0749
OC705	Dif A	0.05363	19.15	0.1366	-37.07	38.02
CD5501	Dif B	0.4022	3.69	0.02836	-6.787	7.704
E2	Dif C	0.1702	35.33	0.2457	-68.97	70.08
OC705	Dif A	-0.2092	21.22	0.1481	-41.98	41.49
E1	Dif B	6.32	2.784	0.01988	0.8648	11.83
E2	Dif C	0.1298	33.46	0.2557	-64.82	64.93
CD5501	Dif A	0.02703	18.06	0.1277	-35.86	35.8
E1	Dif B	3.248	3.144	0.02068	-2.937	9.391
E2	Dif C	-0.0226	33.23	0.2489	-65.3	63.9
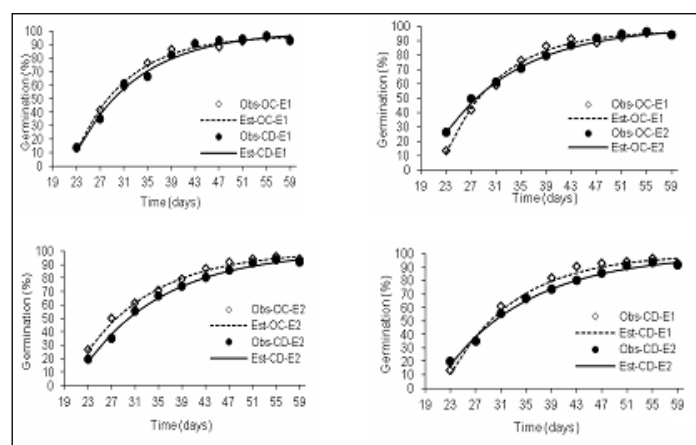
Fig 2. Observed (Obs) and estimated (Est) seed germination curves for the OC705 and CD5501 hybrids sown on two different dates (E1, 28 October; and E2, 20 November).

Table 3. Time (days) for maximum germination percentage and highest dry biomass accumulation of the OC705 and CD5501 hybrid seeds sown on 28 October (E1) and 20 November (E2).

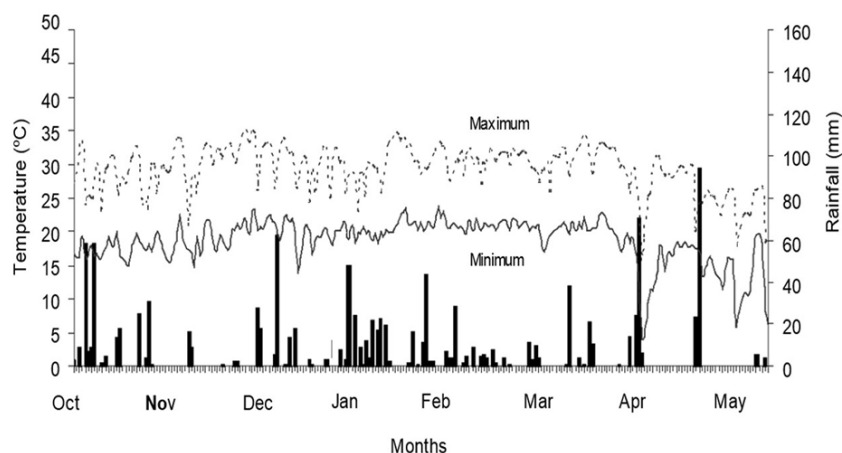| Variable	Epochs	Hybrids	Time (days)	Germination (%)
Germination	October 28	OC705	56	95.0
		CD5501	57	96.2
	November 20	OC705	57	95.3
		CD5501	58	93.3
Biomass	October 28	OC705	48	0.2511
		CD5501	50	0.2203
	November 20	OC705	39	0.1901
		CD5501	40	0.1702

Fig 3. Environmental conditions under which the field tests were conducted.

The accumulated biomass was described by the empirical nonlinear sigmoidal model as follows:

$$y_i = A/(1 + \exp((B - t_i)/C) + \varepsilon_i, \quad i=1, \quad 2, \quad ..., \quad n$$

(2)

Where; y is the dry biomass, A is the maximum value that the response y reaches, and B and C are the scale parameters.

Assuming that the residuals of the model described by equation (2) are additive and independent, and normally distributed with zero mean and variance (σ_e^2), the following equation is obtained:

$$(y_i \mid \sigma_e^2) \sim N(A/(1 + \exp((B - t_i)/C), \sigma_e^2)$$

The following *a priori* distributions were considered for the parameters of the model: A ~ N(0;10^{-2}), B ~ Gamma(10^{-1};10^{-2}), C ~ Gamma(10^{-1};10^{-2}) and τ ~ Gamma(10^{-3};10^{-3}) such that $\sigma_e^2 = \tau_e^{-1}$.

Analyses were performed using WinBUGS software (Lunn et al., 2000), and the samples were generated through Markov chain Monte Carlo (MCMC) totaling 50,000 iterations. The inferences were performed considering the last 40,000 iterations with a jump size of 50 to obtain convergent samples to be used in the inference of the parameters of the *a posteriori* marginal distributions.

To test the hypothesis of equality of parameters by analyzing the credibility interval, the samples of the *a posteriori* marginal distributions was used for the parameter differences, i.e., the difference was considered a function. If the credibility interval contains the value zero, it is concluded that, on average, the parameters of the two involved populations are equal. If the equality is verified for all parameters, it is concluded that all equations are equal (Filho et al., 2008). The convergence of the chains was analyzed by the same software through graphical tests and the "GRdiag" Gelmam-Rubin diagram.

Conclusion

The Bayesian study described the percentage of seed germination and the accumulated dry biomass during seed maturation. Hybrids in two sowing dates were compared using credible intervals. The first sowing date was the most favorable to plant the two hybrids that must be collected from 57 days after the female flowering period when it reaches the maximum germination percentage. The highest cumulative dry biomass was 48 days for hybrid OC705 the E1 time. The results indicate that the Bayesian inference is useful in studying the germination percentage and the dry biomass accumulation should be used in seed area.

References

Andreoli C, Andrade RV (2007) Seed longevity chart to predict viability of corn seed during open storage. Revista Brasileira de Milho e Sorgo. 6: 247-255.

Araujo EF, Araujo RF, Sofiatti V, Silva RF(2006a) Maturação de sementes de milho-doce – grupo Super Doce.RevBras Sementes. 28: 69-76.

Araujo EF, Araujo RF, Sofiatti V, Silva RF (2006b) Qualidade fisiológica de sementes de milho-doce colhidas em diferentes épocas. Bragantia. 65: 687-692.

Balestre M, Pinho RGV, Souza Junior CL, Bueno Filho JSS (2012a) Bayesian mapping of multiple traits in maize: the importance of pleiotropic effects in studying the inheritance of quantitative traits. Theor Appl Genet. 125: 479-493.

Balestre M, Pinho RGV, Brito AH (2012b) Bayesian inference to study genetic control of resistance to gray leaf spot in maize. Genet Mol Res. 11: 17-29.

Borba CS, Andrade RV, Azevedo JT, Oliveira AC (1994) Maturidade fisiológica de sementes do híbrido simples BR 201 de milho (*Zeamays* L.). RevBras Sementes. 16: 63-67.

Borba CS, Andrade RV, Azevedo JT (1995) Maturidade fisiológica de sementes do híbrido simples BR 201 fêmea de milho (Zeamays L.) produzidas no inverno. Rev Bras Sementes. 17: 129-132.

Brasil. Ministério da Agricultura, do Meio-Ambiente e da Reforma Agrária. (1992), Regras para Análises de Sementes. Brasília: Departamento Nacional de Produção Vegetal.

Cané-Retamales C, Mora F, Vargas-Reeve F, Perret S, Contreras-Soto R (2011) Bayesian threshold analysis of breeding values, genetic correlation and heritability of flowering intensity in *Eucalyptus cladocalyx* under arid conditions. Euphytica. 178:177-183.

David AMSS, Araujo EF, Miranda GV, Dias DCFS, Galvão JCC, Carneiro V (2003)Maturação de sementes de milho-pipoca. Revista Brasileira de Milho e Sorgo. 2: 121-131.

Filho SM, Silva FF, Carneiro APS, Muniz JA (2008) Abordagem Bayesiana das curvas de crescimento de duas cultivares de feijoeiro. Cienc Rural. 38: 1516-1521.

Geweke J (1992) Evaluating the accuracy of sampling based approaches to the calculation of posterior moments. In: Bernardo JM, Berger JO, David AP, Smith AFM (eds.) Bayesian statistics. Oxford University, New York, NY, USA.

Kass RE, Raftery AE (1995) Bayes factors. J Am Stat Assoc. 90: 773-795.

Lunn DJ, Thomas A, Best N, Spiegelhalter D (2000) WinBUGS - a Bayesian modelling framework: concepts, structure, and extensibility. Stat Comput. 10: 325-337.

Marcondes MC, Andreoli C, Miglioranza E (2011) Viability equation to determine the longevity of fungicide-treated seeds of wheat stored in a conventional warehouse. ActaSci-Agron. 33: 539-544.

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics. 157, 1819-1829

Mora F, Gleadow R, Perret S, Scapim CA (2009) Genetic variation for early flowering, survival and growth in sugar gum (*Eucalyptus cladocalyx* F. Muell) in southern Atacama Desert. Euphytica. 169:335-344.

Mora F, Serra N (2014) Bayesian estimation of genetic parameters for growth, stem straightness, and survival in Eucalyptus globulus on an Andean Foothill site. Tree Genet Genomes. 10:711-719

Ratkowsky DA (1983) Nonlinear regression modeling. Marcel Dekker, New York, NY, USA.

Sorensen D, Gianola D (2002) Likelihood, Bayesian and MCMC methods in quantitative genetics. Springer, Amsterdam.

Saavedra J, Silva TA, Mora F, Scapim CA (2013) Bayesian analysis of the genetic structure of a Brazilian popcorn germplasm using data from simple sequence repeats (SSR). Chil J Agr Res. 73: 99-107.

Safner T, Miller MP, McRae BH, Fortin MJ, Manel S (2011) Comparison of Bayesian clustering and edge detection methods for inferring boundaries in landscape genetics. Int J Mol Sci. 12:865-889.

Sánchez-Coronado ME, Coates R, Castro-Colina L, Buen AG, Paez-Valencia J, Barradas VL, Huante P, Orozco-Segovia A (2007) Improving Seed Germination and Seedling Growth of *Omphaleaoleifera* (Euphorbiaceae) for Restoration Projects in Tropical Rain Forests. Forest Ecol Manag. 243: 144-155.

Silva FF, Viana JMS, Faria VR, Resende MDV (2013) Bayesian inference of mixed models in quantitative genetics of crop species. Theor Appl Genet. 126: 1749-1761.