SELECTION OF *Coffea arabica* L. HYBRIDS USING MIXED MODELS WITH DIFFERENT STRUCTURES OF VARIANCE-COVARIANCE MATRICES

Fernanda Aparecida Castro Pereira[1], Samuel Pereira de Carvalho[2], Tiago Teruel Rezende[3],
Leonardo Luiz Oliveira[4], Diego Rosa Baquião Maia[5]

ABSTRACT: This study aimed to evaluate different structures of variance-covariance matrices in modeling of productive performance of coffee genotypes over the years, and select hybrids of *Coffea arabica* using mixed models. A mixed linear model was used to estimate variance components, heritability coefficients, and prediction of genetic values of hybrids and cultivars. Three commercial cultivars and eight hybrids of *C. arabica* L. were evaluated. The field production after acclimatization of seedlings was conducted in March 2006. The yield averages from 2009, 2010, 2011, 2013, and 2014 agricultural years were evaluated. The selection criteria of models were used to test 10 structures of variance-covariance matrices, and later a model was chosen to estimate the components of variance, heritability coefficients, and prediction of genetic values. According to Bayesian information criterion (BIC), the best structure was ARMA (Autoregressive Moving Average); however, considering the Akaike Information Criterion (AIC) and corrected Akaike Information Criterion (AICC), the CSH (Heterogeneous Composite Symmetric) was indicated. The Spearman correlation between the genotypic values obtained in the models with ARMA and CSH type R matrix was 0.84. The high and positive correlation indicates that the best model could involve the R matrix with ARMA or CSH structure. The heritability of individual genotypes differed from heritability in broad sense, which considers the independence among agricultural years. Hybrids with higher performance were identified by ordering the genotypic effects, among them, H 2.2, H 4.2, and H 6.1 hybrids were highlighted.

Index terms: Plant breeding, repeated measures, yield.

SELEÇÃO DE HÍBRIDOS DE *Coffea arabica* L. POR MEIO DE MODELOS MISTOS COM DIFERENTES ESTRUTURAS DE MATRIZES DE VARIÂNCIA E COVARIÂNCIA

RESUMO: Neste trabalho buscou-se avaliar diferentes estruturas de matrizes de variâncias e covariâncias na modelagem do comportamento produtivo de genótipos de cafeeiro ao longo dos anos, e selecionar híbridos de *Coffea arabica* utilizando um modelo linear misto Foram avaliadas três cultivares comerciais e oito híbridos de *C. arabica* L., utilizando mudas clonais obtidas por meio do enraizamento de estacas caulinares de ramos ortotrópicos. O plantio em campo, após aclimatação das mudas, foi realizado em março de 2006. Foram avaliadas as produções médias dos anos agrícolas 2009, 2010, 2011, 2013 e 2014. Critérios de seleção de modelos (RLL, AIC e BIC) foram utilizados para testar 10 estruturas da matriz R de variâncias e covariâncias, e posteriormente selecionou-se um modelo para estimar os componentes de variância, os coeficientes de herdabilidades e predição dos valores genéticos. Segundo o critério de informação Bayesiano (BIC) a melhor estrutura foi a ARMA, porém considerando o critério de Akaike e Akaike corrigido, a CSH foi indicada. A correlação de Spearman entre os valores genotípicos obtidos nos modelos com matriz R do tipo ARMA e CSH foi 0,84, alta e positiva, indicando que nesse conjunto de dados o melhor modelo pode envolver a matriz R com estrutura ARMA ou CSH. A herdabilidade individual dos genótipos foi diferente da herdabilidade no sentido amplo, que considera a independência entre os anos agrícolas. Pelo ordenamento dos efeitos genotípicos foram identificados os híbridos de desempenho superior, dentre os quais destacaram os híbridos H 2.2, H 4.2 e H 6.1.

Termos para indexação: Melhoramento genético, medidas repetidas, produtividade.

1 INTRODUCTION

The importance of coffee agribusiness can be evaluated not only by production and profits, but also by its role in the job market as a generator of jobs and as a factor of fixation of labor in the rural environment (SANTOS et al., 2009). Given this outlook and the importance of coffee production chain, coffee breeding programs have an important role in growing coffee in Brazil and become more economically competitive. The breeding of arabic coffee is more directed to development of pure-lines cultivars (MEDINA-FILHO et al., 2008), however, there is a great potential for use of hybrids with high heterosis (BERTRAND et al., 2011; MOHHAMED, 2011) and with multiple resistance to important diseases (ANDREAZI et al., 2015).

[1]Universidade Federal de Lavras/UFLA - Departamento de Biologia/DBI - 37.200-000 - Lavras - MG - fernandacpereira01@gmail.com
[2]Universidade Federal de Lavras/UFLA - Departamento de Agricultura/DAG - 37.200-000 - Lavras -MG - samuelpc@dag.ufla.br
[3]Universidade José do Rosário Vellano - UNIFENAS - Rod. MG 179, Km 0 - 37.132-440 - Alfenas- MG - tiago.rezende@unifenas.br
[4]Secretaria de Agricultura e Abastecimento de Cristais - Praça Cel. Joaquim Luiz da Costa Maia, 01 - Centro - 37.275-000 - Cristais - MG leonardooliveira04@outlook.com.br
[5]Rehagro Consultoria Agronômica Rua Santa Fé, 100, Sion - 30.320-130 - Belo Horizonte - MG - diego.baquiao@rehagro.com.br

Perennial plant species such as coffee trees have peculiar aspects, including a long reproductive cycle, marked annual oscillation of production resulting in a biennial cycle, overlap of generations, expression of characters over several years, and differences in precocity and productive longevity (BRUCKNER, 2008). Due to these agronomic peculiarities, coffee breeding is difficult (OLIVEIRA et al., 2011). Therefore, the use of special methods to estimate genetic parameters and predict genetic values is recommended (DE FAVERI et al., 2016).

When researchers deal with segregating coffee populations, the selection on a single plant level is relevant and is not only based on the average progeny. Thus, special methods to estimate the genetic parameters and prediction of genotypic values are necessary (PEREIRA et al., 2013; RESENDE, 2007). The application of mixed models in plant breeding has emphasized the estimation of variance components and appropriate identification of experimental error to test the hypotheses of fixed effects (HENDERSON, 1975). A general proposal has rarely been used when considering the modeling of variance-covariance genetic structures and random effects predictions (PIEPHO et al., 2008). Some studies have detected the small changes in the parameter estimation as a function of the variance-covariance structures used (APIOLAZA; GARRICK, 2001), including a change in genotype ordering when applying the selection to different models with various structures (ANDRADE et al., 2016).

The objective of this work was to evaluate different matrix structures of variance-covariance in the modeling of the yield performance of coffee genotypes over the years. A mixed linear model was used for the estimation of variance components, heritability coefficients, and genetic values prediction of hybrids and existing cultivars.

2 MATERIAL AND METHODS

The experiments were conducted at the Department of Agriculture, Coffee Sector at the Federal University of Lavras (UFLA). Three cultivars and eight hybrids were evaluated, as described in Table 1.

The seedlings of hybrids were obtained by rooting stem cuttings of orthotropic branches harvested from matrices plants. Segments 3–5 cm in length were treated with phytohormone indole-3-butyric acid (IBA) in inert talc and placed into commercial substrate to rooting. Rooting

was conducted in greenhouse that had humidity and temperature control and was equipped with automatic irrigation system. The substrate used for the rooting was a mixture of washed sand and vermiculite [1:1]. After rooting, cuttings were transplanted to conventional polyethylene sacks (10×20 cm) for half-year seedlings. The sacks contained Plantmax®, a commercial substrate and standard substrate in a 1:1 ratio. Seedlings were then transferred to a nursery under sunlight with 50% of shading, where they remained until they reached the recommended seedling size for field planting.

After acclimatization, seedlings were planted in March 2006, following recommendations of planting (typical cultural practices) for the region. The average grain yield of years 2009, 2010, 2011, 2013, and 2014 was evaluated using mixed model analyses. The following matrix model was considered assuming normal multivariate distribution:

$$Y = X_r + Z_g + W_a + ZW_{ga} + \varepsilon$$

$$Where \begin{cases} y|b, V \sim N\,(Xb, V) \\ g \sim N\,(0, I\sigma_g^2) \\ a \sim N(0, I\sigma_a^2) \\ ga \sim N(0, I\sigma_{ga}^2) \\ \varepsilon \sim N\,(0, R) \end{cases}$$

In this model, Y is the data vector, r is the vector of blocks effects (assumed to be fixed) added to the general mean, g is the vector of individual genotypic effects (assumed to be random), a is the vector of effects of agricultural years (random), ga is the vector of the interaction effects of genotypes and agricultural years. In the previous description, the random error vector associated with the experimental unit is distributed as N (O,R), that is, normal with mean 0 and a positive and defined dimension covariance matrix R. In general, crop data are analyzed together by assuming a composite symmetry (CS) of matrix R, with equal variances and null correlations, or experimental design considering model as split-plots in time. However, this kind of assumption may not be accurate. To verify the validity of this assumption, sphericity hypothesis of residual covariance matrix was applied by Mauckhly sphericity test (MAUCKHLY, 1940) using SAS® University version (PROC GLM).

TABLE 1 - Identification and description of the materials (C: cultivars; H: hybrids) evaluated in the trial.

Identification	Description
C 1	Icatu IAC-2942
C 2	Catuaí IAC-62
C 3	Catuaí IAC-99
H 1.2	(Icatu IAC-2942 x Catuaí IAC-62); Plant 2
H 1.3	(Icatu IAC-2942x Catuaí IAC-62); Plant 3
H 2.1	(Icatu IAC-2942x Icatu IAC-5002); Plant 1
H 2.2	(Icatu IAC-2942x Icatu IAC-5002); Plant 2
H 4.1	(Icatu IAC-4040-179 x Catuaí IAC-17); Plant 1
H 4.2	(Icatu IAC-4040-179 x Catuaí IAC-17); Plant 2
H 6.1	(Icatu IAC-4040-179 x Catuaí IAC-99); Plant 1
H 6.2	(Icatu IAC-4040-179 x Catuaí IAC-99); Plant 2

After sphericity test, 10 alternative structures of R matrix were tested in order to deal with present reality, to better explore data, and thus, obtain accurate estimates. The structures tested were: Antidependence (ANTE), Autoregressive (AR), Autoregressive with heterogeneous variances (ARH), Autoregressive moving averages (ARMA), Composite symmetry with heterogeneous variances (CSH), Analytical Factor (FA), Huynh-Feldt (HF), Toeplitz (TOEP), Toeplitz with Heterogeneous Variance (TOEPH), and Unstructured (UN).

The parameters related to variances of each evaluation are generally located within the main diagonal of matrices. Outside of diagonal are covariance parameters for each year pair considered. The number of parameters estimated by each covariance structure is a function of global dimension (t) of covariance matrix. For matrix R modeling, the (t) variable corresponds to the number of harvests (Table 2).

Three criteria were considered when selecting models under different covariance matrix structures: 1) Maximum likelihood ratio (ML) test, 2) Akaike model selection criteria (AIC), and 3) Schwarz criterion (BIC) (KONISH and KITAGAWA, 2008; GURKA, 2006). These criteria indicate which model is the most likely among analyzed ones and referring to data in question and do not guarantee the choice of the true model.

The likelihood ratio testing evaluates whether the additional parameters significantly improve the model. This test considered $L1 = -2\log(L)$ as the model with the lowest

number of parameters and $L2 = -2\log(L)$ as the model with the highest number of parameters. The tested hypothesis was whether the two models were equivalent (i.e., the extra parameters do not differ from zero). Under normality, the difference between L1 and L2 is asymptotically distributed as L1 - L2 ~ χ2 [r] (chi-square with r degrees of freedom).

As an alternative to maximum likelihood ratio test, it is possible to calculate measures based on information such as AIC and BIC criteria. The information is calculated as a penalty term applied to the likelihood function. The AIC criterion is based on decision theory and in order to avoid excessive parameterization, penalizes models with large number of parameters. It is defined by expression:

$$AIC = -2\log(L) + 2p$$

Where log (L) is the logarithm of the maximum likelihood function of the model, and p is the number of parameters of the variance-covariance matrix. According to this criterion, the variance-covariance matrix model to be chosen is the one with the lowest AIC value. The Schwarz criterion or Bayesian information criterion (BIC) was derived from the Bayes' theorem to problem of model identification. The BIC value is minimized asymptotically in order of the model with the highest probability later. It is defined by expression:

$$BIC = -2\log(L) + \log(N) * p$$

TABLE 2 - Description of the tested variance-covariance structures, and the number of parameters tested in each structure.

Structures	Description	Parameters*
ANTE(1)	Ante-Dependence	2t-1
AR(1)	Autoregressive	2
ARH(1)	Heterogeneous Autoregressive	t+1
ARMA(1,1)	Autoregressive Moving Average	3
CSH	Composite Symmetric with Heterogeneous variance	t+1
FA(1)	Factor Analytic	q/2(2t-q+1)+t
HF	Huynh-Feldt	t+1
TOEP	Toeplitz	t
TOEPH	Heterogeneous Toeplitz	2t-1
UN	Unstructured	t(t+1)/2

*t is number of harvests

Where log (L) is the logarithm of the maximum likelihood function of the model, N is the total number of observations and p is the number of parameters of the variance-covariance matrix. While ML, AIC, and BIC are conceptually different, they use the same statistical criterion: the maximum of the likelihood which functions as a measure of adjustment. However, these criteria define different critical values (LITTELL et al., 2006).

After selection of the best structure of variance-covariance matrix, it was possible to obtain variance estimates by restricted maximum likelihood (REML) method and to predict the genetic values by Empirical Best Linear Unbiased Prediction (E-BLUPs) using SAS University Version (PROC MIXED). The E-BLUP has good predictive accuracy compared to other procedures (PIEPHO et al., 2008); a common estimator used was based on E-BLUP expression, that is:

$$E-BLUP = h^2(\bar{Y}_{i..} -$$

(BERNARDO, 2010). Therefore, it was used to calculate the heritability of each genotype by expression:

$$h_i^2 = \frac{E-BLUP}{(\bar{Y}_{i..} - \bar{Y}_{...})}$$

Where $\bar{Y}_{i..}$ is the genotype i mean in all harvests, and $\bar{Y}_{...}$ is the overall mean of all genotypes in all harvests under consideration. And broad sense heritability considering the composite symmetry according to Ramalho et al. (2012):

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_{ga}^2}{a} + \frac{\sigma_r^2}{ga}}$$

Where σ_g^2 is the variance of the genotypes; σ_{ga}^2 is the variance of the interaction genotypes and agricultural years; σ_r^2 is the residual variance; a is the number of agricultural years and g is the number of genotypes.

A coincidence index (CI) of the top three selected genotypes was obtained according to Hamblin and Zimmerman (1986):

$$CI = \frac{A - C}{B - C}$$

Where, A is number of coincident progenies among the top three genotypes selected; B is the number of selected progenies. C is the expected amount of coincident progenies.

3 RESULTS AND DISCUSSION

Before comparing the quality of adjustment associated with the different models, the correlations between the years were estimated according to the means of all genotypes in all replicates (Table 3).

The highest positive correlation was identified between the years 2011 and 2014 (r = 0.41, $p < 0.05$). Negative correlations were identified in subsequent years (2010 and 2011; 2013 and 2014).

TABLE 3 - Pearson's correlations among the evaluation years regarding genotype yield (hybrids and cultivars).

Years	2009	2010	2011	2013
2010	0.304			
2011	0.220	-0.16		
2013	0.131	0.030	-0.012	
2014	0.0192	0.22	0.412	-0.400

Therefore, it was demonstrated that the assumption of independence among years cannot be accepted as a rule to support a model for traditional variance analysis (FREITAS et al., 2008). This fact is confirmed by Mauchly's sphericity test (Table 4).

The hypothesis of sphericity was rejected by the Mauchly's test, which indicated that the composite symmetry structure (CS) for the variance-covariance matrix might not be the most adequate for the data in question. According to Resende (2007), the sphericity rejection in perennial plants has been very common. With the sphericity rejection, 10 alternative structures were examined for the covariance matrices associated with matrix R.

The answers of the different criteria of information and selection of models diversify according to some characteristics—such as the probability distribution of the data—and, primarily, the covariance pattern present in such characteristics. The AIC and BIC criteria are RLL adjustments and they are the most used in the literature (GURKA, 2006). A lower value of these statistics indicates a better structure. At the 5% probability level, the structure $I\sigma^2$ (used in PROC GLM) is not the most appropriate compared to the structures tested. Thus, according to BIC, the best structure was ARMA. This brings together characteristics of first order autoregressive structure with the moving average technique (COOPER and THOMPSON, 1977). However, when considering AIC and AICC, the CSH structure is the best option (Table 5). In general, the two criteria produce concordant results (FLORIANO, 2006). However, in this way, the choice must account for number of parameters.

The use of these techniques is fundamental to decision theory in mixed models; besides the quality of adjustment, such techniques consider the principle of parsimony, which penalizes models with a greater number of parameters (BURNHAM and ANDERSON, 2004).

For perennial and semi-perennial crops, trials of which are usually evaluated in different years, there is a case of repeated measures in time (PIEPHO and ECKL, 2014). The correlation between years may decrease with the temporal distance between harvests (SONG, 2007). One reason is that genes expressed in the first agricultural year may not be expressed in subsequent years (PASTINA, 2012).

The best structure regarding the ability of the model to fit the observed data varies between trials, revealing the impossibility of previously indicating a structure for the analysis of similar experiments (SILVA, DUARTE, REIS, 2015). For example, a similar study examining *C. arabica*, Andrade et al. (2016) identified that the Toeplitz (TOEP) structure for the variance-covariance matrix is the most appropriate because it considers specific correlations for each interval between harvests. In addition, Burgueño et al. (2011) demonstrated that the predictive power of a model increased up to 6% when the variance-covariance structure was adequately modeled.

Table 6 presents the genotypic values of genotypes under evaluation considering the two primary variance and covariance structures.

The cultivars presented low genotypic values when compared to hybrids. The order of classification of genotypic values for ARMA and CSH were coincident just in first two positions, with hybrids H 2.2 and H 4.2 and in last position with C3. However, by Spearman's correlation, there was a high (0.82) and significant (P-value = 0.002) correlation between the E-BLUPS of two structures in question. In practical terms, in selection of three best genotypes, the coincidence index was 66%. However, when the four best genotypes were selected, the coincidence index was 75%. These values are high due to low population sampled. However, in larger populations, it is possible that the coincidence percentage decreases, requiring even more apropriate choice of the best structure of variances-covariance matrix.

To calculate the heritability of the progenies, the *R* matrix of the ARMA type was considered. As a result, a low heritability of the genotype (12.3%) was found in comparison to the heritability in the broad sense (73.5%). This fact is an alert for the mild selection of genotypes. The genetic variance was lower than the environmental variance, due to the field experimentation conditions that tended to increase environmental variance.

TABLE 4 - P-values of the Mauchly's sphericity test according to coffee yield evaluated in the years 2009, 2010, 2011, 2013, and 2014.

Degrees of freedom	Mauchly's criterion(W)	X^2	p-value
9	0.2136	26.8851	0.0015

TABLE 5 - Results of the tests for the best *R* matrix choice for the model involving coffee yield evaluated in years 2009, 2010, 2011, 2013, and 2014.

Structures	F Qui	p-value	ML[1]	AIC[2]	AICc[3]	BIC[4]
ANTE(1)	4 10.14	0.0381	1476.4	1486.4	1486.8	1496.5
AR(1)	1 3.87	0.0491	1482.7	1486.7	1486.8	1490.7
ARH(1)	3 10.14	0.0174	1476.4	1484.4	1484.7	1492.5
ARMA(1,1)	1 6.85	0.0089	1479.7	1483.7	1483.8	1487.7
CSH	3 13.28	0.0041	1473.3	1481.3	1481.5	1489.3
FA(1)	5 13.60	0.0183	1473.0	1485.0	1485.5	1497.0
HF	3 11.62	0.0088	1474.9	1482.9	1483.2	1491.0
TOEP	2 6.99	0.0304	1479.6	1485.6	1485.7	1491.6
TOEPH	4 13.60	0.0087	1473.0	1483.0	1483.4	1493.0
UN	5 13.60	0.0183	1473.0	1483.0	1483.4	1493.0

1: Maximum Likelihood Test

2: Akaike Information Criterion

3: Akaike Information Criterion Corrected

4:BayesianInformationCriterion(BIC)

TABLE 6 - Genotypic values of the coffee genotypes (C: cultivars, H: hybrids) in terms of yield considering the years 2009, 2010, 2011, 2013 and 2014.

Identification of coffee genotypes*	Structures of *R* matrix	
	ARMA(1,1) Autoregressive Moving Average	CSH Heterogeneous Composite Symmetric
C 1	0.20	0.01
C 2	-0.66	-0.02
C 3	-4.26	-0.14
H 1.2	-2.57	0.13
H 1.3	1.38	0.50
H 2.1	-1.82	-0.07
H 2.2	9.39	9.46
H 4.1	3.68	0.12
H 4.2	5.81	7.68
H 6.1	1.60	2.51
H 6.2	1.54	2.42

* Description in Table 1

4 CONCLUSIONS

The best variance-covariance structures for the data analysis of the experiment in question were ARMA and CSH. The genotypic values obtained by these two structures are correlated, however, when simulating genotype selection, a minimum of four genotypes is expected so that at least three genotypes are selected based on the analyzes with the two structures. Therefore the three best hybrids were H 2.2 [(Ic 2942 x 5002); Plant 2], H 4.2 [(Ic 4040-179 x Ct 17); Plant 2] and H 6.1 [(Ic 4040-179 x Ct 99); Plant 1].

5 ACKNOWLEDGMENTS

6 REFERENCES

ANDRADE, V. T. et al. Statistical modeling implications for coffee progenies selection. **Euphytica,** Wageningen, v. 207, n. 1, p. 177-189, Jan. 2016.

ANDREAZI, E. et al. Desempenho de híbridos F_1 de café arábica com resistência simultânea a ferrugem, mancha aureolada e bicho mineiro. **Coffee Science**, Lavras, v. 10, n. 3, p. 375-382, July/Sept. 2015.

APIOLAZA, L. A.; GARRICK, D. J. Analysis of longitudinal data from progeny tests: some multivariate approaches. **Forest Science**, Bethesda, v. 47, n. 2, p. 129-140, May. 2001.

BERNARDO, R. N. **Breeding for quantitative traits in plants**. 2. ed. Woodbury: Stemma Press, 2010, 390p.

BERTRAND, B. et al. Performance of *Coffea arabica* F1 hybrids in agroforestry and full-sun cropping systems in comparison with American pure line cultivars. **Euphytica**, Wageningen, v. 181, n. 2, p 147-158, Sept. 2011.

BRUCKNER, C. H. **Fundamentos do melhoramento de fruteiras.** 1. ed. Viçosa: Editora UFV, 2008, 202 p.

BURGUEÑO, J. et al. Genomic prediction of breeding values when modeling genotype×environment interaction using pedigree and dense molecular markers. **Crop Science**, Madison, v. 52, n. 2, p. 707-719, Mar/Apr. 2012.

BURNHAM, K. P.; ANDERSON, D. R. Multimodel Inference - Understanding AIC and BIC in model selection. **Sociological Methods and Research**, Cambridge, v. 33, p. 261-304, 2004.

CARVALHO, A. Principles and practice of coffee plant breeding for productivity and quality factors. In: CLARKE, R. J.; MACRAE, R. (Eds.). **Coffee: Agronomy**. London: Elsevier Applied Science, 1988. p. 129–165.

COOPER, D. M.; THOMPSON, R. A note on the estimation of the parameters of the autoregressive-moving average process. **Biometrika**, London, v. 64, n. 3, p. 625-628, Sept/Dec. 1977.

DE FAVERI J. et al. Residual variance–covariance modeling in analysis of multivariate data from variety selection trials. **Journal of Agricultural, Biological, and Environmental Statistics**, v. 22, n. 1, p. 1–22, Jan/Mar. 2016.

FLORIANO, E. P. et al. Ajuste e seleção de modelos tradicionais para série temporal de dados de altura de árvores. **Ciência Florestal**, Santa Maria, v. 16, n. 2, p. 177-199, Apr/June. 2006.

FREITAS, E. G. de. et al. Modelo univariado aplicado a dados longitudinais de cana-de-açúcar. **Revista Brasileira de Biometria**, Lavras, v. 26, p. 93-16, 2008.

GURKA, M. J. Selecting the best linear mixed model under REML. **The American Statistician**, Washington, v. 60, n. 1, p. 19-26, Jan/Mar. 2006.

HENDERSON, C. R. Best linear unbiased estimation and prediction under a selection model. **Biometrics**, Bethesda, v. 31, p. 423-449, 1975.

KONISHI, S.; KITAGAWA, G. **Information criteria and statistical modeling.** 1. ed. New York: Springer, ISBN 0387718869. 2008. 321p.

LITTELL, R. C. et al. **SAS system for mixed models**. 2. ed. Cary: Statistical Analysis System Institute, 2006. 828 p.

MAUCHLY, J. W. Significance test for sphericity of a normal n-variate distribution. **Annals of Mathematical Statistics**, Ann Arbor, v. 11, n. 2, p. 204- 209, Apr/June. 1940.

MEDINA FILHO, H. P.; BORDIGNON, R.; CARVALHO, C. H. S. de. Desenvolvimento de novas cultivares de café arábica. In: CARVALHO, C. H. S. de, (Ed.). **Cultivares de café: origem características e recomendações**. Brasília: Embrapa Café, 2008. p. 79-101.

MOHAMMED, W. Heterosis and combining ability analysis for coffee quality in diallel crosses of diverse coffee (*Coffea arabica* L.) parents in origin. **East African Journal of Sciences**, Haramaya, v. 5, n. 1, p. 12-21, Jan/June. 2011.

OLIVEIRA, A. C. B. et al. Prediction of genetic gains from selection in Arabica coffee progenies. **Crop Breeding and Applied Biotechnology**, Viçosa, v. 2, p. 106-113, 2011.

PASTINA, M. M. et al. A mixed model QTL analysis for sugarcane multiple-harvest-location trial data. **Theoretical and Applied Genetics**, New York, v. 124, n. 5, p. 835-849, Mar. 2012.

PEREIRA, T. B. et al. Eficiência da seleção de progênies de café F4 pela metodologia de modelos mistos (REML/BLUP). **Bragantia**, Campinas, v. 72, n. 3, p. 230-236, July/Sept. 2013.

PIEPHO, H. P.; ECKL, T. Analysis of series of variety trials with perennial crops. **Grass and Forage Science**, Victoria, v. 69, n. 3, p. 431–440, July/Sept. 2014.

PIEPHO, H. P. et al. BLUP for phenotypic selection in plant breeding and variety testing. **Euphytica**, Wageningen, v. 161, n. 1-2, p. 209-228, May. 2008.

RAMALHO, M. A. P. et al. **Aplicações da genética quantitativa no melhoramento de plantas autógamas**. 1. ed. Lavras: Editora UFLA, 2012. 522 p.

RESENDE, M. D. V. **Matemática e estatística na análise de experimentos e no melhoramento genético.** 1. ed. Colombo: Embrapa Florestas, 2007. 435 p.

SANTOS, V. E. dos. et al. Análise do setor de produção e processamento de café em Minas Gerais: uma abordagem matriz insumo-produto. **Revista de Economia e Sociologia Rural**, Brasília, v. 47, n. 2, p. 363–388, jun. Apr/June. 2009.

SILVA, E. N.; DUARTE, J. B.; REIS, A. J. dos S. Seleção da matriz de variância-covariância residual na análise de ensaios varietais com medidas repetidas em cana-de-açúcar. **Ciência Rural**, Santa Maria, v. 45, n. 6, p. 993-999, June. 2015.

SONG, P. X. -K. **Correlated data analysis: modeling, analytics, and applications**. New York: Springer. 2007. 352 p.