

Repeatability estimates in longitudinal data on guava trees

Flavia Alves da Silva^{1*}, Alexandre Pio Viana¹, Caio Cezar Guedes Corrêa¹, Lucas Souza da Silva Leal¹, Leonardo Siqueira Glória²

¹Universidade Estadual do Norte Fluminense Darcy Ribeiro/ CCTA – Lab. de Melhoramento Genético Vegetal, Av. Alberto Lamego, 2000 – 28013-602 – Campos dos Goytacazes, RJ – Brasil.

²Universidade Estadual do Norte Fluminense Darcy Ribeiro/ CCTA – Lab. de Ciência Animal, Av. Alberto Lamego, 2000 – 28013-602 – Campos dos Goytacazes, RJ – Brasil.

*Corresponding author <flavia_uems@hotmail.com>

Edited by: Luiz Alexandre Peternelli

Received March 15, 2022

Accepted August 08, 2022

ABSTRACT: The use of longitudinal measurements is an essential practice both in *Psidium guajava* L. breeding and in other perennial crops in which covariance structures can be introduced to explain the form of dependence between measurements. Hence, this study aimed to analyze six covariance structures to identify one that best described the correlation between the repeated measurements in time in traits of guava full-sib families. The repeatability coefficient for each trait was estimated and the minimum number of evaluations required for estimates representing the population was determined. The work was performed based on average data of three yield-related variables from nine harvests of a guava tree population evaluated from 2011 to 2018. The best model was chosen based on the Akaike and Schwarz Bayesian information criterion. The autoregressive covariance structure best represented the dependencies among families between crops for all traits. The number of variables of fruits and total yield per plant presented repeatability estimates higher than 0.5 and may be essential traits for indirect selection of others, such as fruit mass, which had an estimated repeatability of 0.24, proving low regularity in the repetition of the character from one cycle to another. It was also possible to define four harvests as the minimum acceptable number of observations necessary on the same individual for these traits; therefore, the repetitions represented the individuals.

Keywords: *Psidium guajava* breeding, covariance structure, repeated measurements

Introduction

Long yield cycles in perennial plants require repeated measurements of individuals throughout time to estimate variance components with greater accuracy (Mathew et al., 2018). In perennial species, models should consider an additional effect, the well-known permanent environment effect, and phenotypic correlation among repeated measurements in the same individual, known as repeatability (Resende et al., 2006; Resende et al., 2017). The repeatability coefficient measures the capacity of individuals to repeat the trait expression throughout some yield cycles. This parameter is important to predict genotypic values, selective efficiency, and heritability in a minimum of measurements taken for a certain trait (Maia et al., 2013).

The mixed linear models could be used to describe longitudinal data, choosing different matrices for covariance structures associated with random factors of the model that explain the dependence between measurements (Shalizi and Isik, 2019). The simplest repeatability model assumes the independent residual effect and considers environmental and genetic correlations constant among different registers of repeated measurements. Although this may not be a realistic assumption, it is commonly used, resulting in biased estimates of variance components (Mathew et al., 2018).

Covariance structures describe different patterns of dependence, ranging from standard repeatability with few parameters but with constant covariances to hyper-parametrized models that lead to overfitting and are computationally infeasible (Wade and Quaas, 1993;

Wolfinger, 1993). However, no structure fits well in all populations of perennial plants, including guava. In this sense, this study aimed to analyze six covariance structures to identify the structure that best described the correlation between the repeated measurements for traits in guava full-sibs. Additionally, the repeatability coefficient for each trait was estimated and the minimum number of evaluations needed for estimates representing the population was determined.

Materials and Methods

We used guava tree families (*Psidium guajava* L.) from established crosses based on genetic diversity. The population is part of the final experiments of a guava tree genetic breeding program before the trials of growing value and use. Harvesting began after the end of the juvenile period of the plants, following the cycle of phytosanitary treatments: intermittent plant period, yield pruning, fertilization, and yield.

The experiment comprised a randomized block design, with two replicates, 17 segregating families, and 12 plants per family, evaluated during nine harvests. Three traits were evaluated in terms of the individual: fruit mass in g (FM), total number of fruits (NF), and total yield per plant in g (TY).

The procedure suggested by Littell et al. (2006) for mixed model analysis was adopted. Firstly, covariance structures were chosen. Subsequently, the fixed effects were specified, followed by the choice/estimate of the covariance structure. After that, the effects of treatment and time were evaluated using generalized minimum squares with the covariance estimated and after the statistical inference based on the results was conducted.

Using the SAS software, the model was adjusted for each covariance structure at a time using the REPEATED statement in the PROC MIXED procedure (Littell et al., 2006). The restricted maximum likelihood was used as an estimator (Patterson and Thompson, 1971) in the model:

$$Y_{ijklm} = \mu + P_i + F_{ij} + B_{ijk} + M_l + e_{ijkl} \quad (1)$$

where: Y_{ijkl} denotes the measurement in the l^{th} harvest in the k^{th} block in the j^{th} family of the i^{th} plant. $\mu + P_i + F_{ij} + B_{ijk} + M_l$ is the mean of the plant i within the family j of block k in the harvest l , containing the effects of family, block, and harvest, respectively. e_{ijkl} is the random error associated with the measurement in the harvest l in the i^{th} plant associated with j^{th} family of the k^{th} block, with $\sim NID(0, R)$.

The distinctive characteristic of a repeated measurement model is the variance and covariance structure of the error e_{ijkl} . Although the plants were randomly attributed to the families, which were randomly attributed to the blocks, the repeated measurement factor levels, the time in this case, are not randomly attributed to the units in the plants. The random errors e_{ijkl} for the same plant are thus not independent. Rather, it was assumed that errors for different plants are independent:

$$Cov[e_{ijk'l}, e_{ijk'l'}] = 0 \quad (2)$$

If $i \neq i', j \neq j',$ or $k \neq k'$.

Moreover, as measurements in the same plant last for a period, they can have different variations and the correlations between pairs of measurements depend on the length of the time interval between measurements. Hence, in general, it was assumed

$$Var[e_{ijkl}] = \sigma_l^2 \quad (3)$$

and

$$Cov[e_{ijkl}, e_{ijk'l'}] = \sigma_{ll'} \quad (4)$$

It was allowed that the variation of e_{ijkl} depended on the time of measurement/harvest l , and the covariance between errors in two harvests, l and l' , for the same plant, depended on the harvests. The covariance model can be expressed according to some structures involving fewer parameters in R . The following covariance structures of the errors were evaluated:

Compound symmetry (CS), characterized by equality of variance and covariance:

$$\sigma^2 \begin{bmatrix} 1.0 & \rho & \dots & \rho \\ & 1.0 & \dots & \rho \\ & & \ddots & \vdots \\ & & & 1.0 \end{bmatrix} = \begin{bmatrix} \sigma^2 + \sigma_1 & \sigma_1 & \dots & \sigma_1 \\ & \sigma^2 + \sigma_1 & \dots & \sigma_1 \\ & & \ddots & \vdots \\ & & & \sigma_1 \end{bmatrix}$$

First-order autoregressive (AR), characterized by equality of variance, and covariance decreases as distances between harvests increase:

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & \dots & \rho^{n-1} \\ & 1 & \dots & \rho^{n-2} \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}$$

Variance component (VC), with homogeneous variance of harvests and absent covariance:

$$\begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ & \sigma_1^2 & \dots & 0 \\ & & \ddots & \vdots \\ & & & \sigma_l^2 \end{bmatrix}$$

Heterogeneous first-order autoregressive (HAR), heterogeneous variances and the covariance between two adjacent measurements is equal to the correlation (r), and the covariance between two non-adjacent measurements is the correlation raised to the number of measurements between the two -1:

$$\begin{bmatrix} \sigma_1^2 & \sigma_2 \sigma_1 \rho & \dots & \sigma_4 \sigma_1 \rho^{n-1} \\ & \sigma_2^2 & \dots & \sigma_4 \sigma_2 \rho^{n-2} \\ & & \ddots & \vdots \\ & & & \sigma_n^2 \end{bmatrix}$$

Compound symmetry with heterogeneous variance (HCS), characterized by inequality of variances:

$$\begin{bmatrix} \sigma_1^2 & \sigma_1 \sigma_2 \rho & \sigma_1 \sigma_3 \rho & \sigma_1 \sigma_n \rho \\ & \sigma_2^2 & \dots & \sigma_1 \sigma_n \rho \\ & & \ddots & \vdots \\ & & & \sigma_n^2 \end{bmatrix}$$

Unstructured (UN), using different variances for each of the l^{th} harvests and different covariances between measurements in different harvests:

$$\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{1n} \\ & \sigma_2^2 & \sigma_{23} & \sigma_{2n} \\ & & \ddots & \vdots \\ & & & \sigma_n^2 \end{bmatrix}$$

Two model adjustment measurements were obtained for each model. The first one was the Akaike Information Criterion (AIC) (Akaike, 1974):

$$AIC = -2 \log L(\hat{\theta}) + 2d \quad (5)$$

where: $\hat{\theta}$ represents the total number of fixed effect parameters and variation components estimated in the

model. The second adjustment measurement of the model was the Bayesian Information Criterion (BIC) (Schwarz, 1978):

$$BIC = -2\log f(x_n|\theta) + p\log n \tag{6}$$

where: $f(x_n|\theta)$ is the model chosen; p is the number of parameters to be estimated; n is the number of observations in the sample.

The accuracy function for permanent phenotypic effects was obtained by:

$$r_{pp}^2 = \frac{mp}{1 + (m - 1)p} \tag{7}$$

where: r_{pp}^2 is the permanent phenotypic accuracy; m , the number of measurements per individuals; and p , the repeatability coefficient.

The efficiency regarding the use of only one harvest was calculated as described by:

$$r_{aa}^2 = \left\{ \frac{mp}{1 + (m - 1)p} \right\}^{1/2} \tag{8}$$

where: r_{aa}^2 is the efficiency in relation to the use of only one harvest; m is the measurements per individuals; and p is the repeatability coefficient.

The coefficient of determination, which represents the prediction certainty of the individual true value for the variables considering the number of measurements performed, was calculated by the equation (Cruz et al., 2012):

$$R^2 = m.p/1 + p.(m - 1) \tag{9}$$

where: R^2 is the coefficient of determination for the number of repetitions made; m is the number of measurements per individuals; and p is the repeatability coefficient.

The estimate of the number of measurements (n_0), required to predict the individual true value with the value of genotype determination (R^2) expected, was determined using the equation (Cruz et al., 2012):

$$n_0 = \frac{R^2(1 - r)}{\{(1 - R)^2\}r} \tag{10}$$

where: R^2 is the coefficient of determination for the number of repetitions made; and r is the number of measurements.

Results and Discussion

The term 'repeated measurement' is used for datasets with several measurements of a response variable in the same experimental unit. In most applications, several measurements are taken during a time. Generally, any data measured repeatedly over time or in space are repeated measurement data. The covariance structure of the data observed is what differentiates data analysis from repeated measurements. In randomized block designs, treatments are randomized to units within a block. It renders all observations within a particular block equally correlated. In repeated measurement experiments, however, two measurements made at adjacent time points are more likely to correlate than two measurements made at multiple time points.

The critical point in these models is a correct specification of the covariance structure to obtain efficient estimates in the analysis of repeated measurements. Six covariance structures were tested in this work. The best model was chosen based on Akaike (AIC) and Schwarz Bayesian (BIC) information criteria for the agronomic performance variables of full-sibs of *Psidium guajava* (Table 1).

No convergence of the iterative process occurred for the models that considered the HAR, HCS, and UN covariance structures, possibly because in each iteration, the residual variation was calculated after the equations of the mixed model were solved and the -2 Log Res Like was obtained. If the difference between -2 Res Log Like of each iteration is less than 1E-8, the model will converge. If the model response continues to vary between iterations, it may not converge (Littell et al., 2006). This result suggests that these structures may not be appropriate for the residue modeling of the set used in this work to obtain repeatability estimates with greater accuracy.

A further reason for the non-conversion of the model is that the covariance matrix can be defined as non-

Table 1 – Values for the AIC (Akaike Information Criterion) and BIC (Schwarz Bayesian) information criteria for the model adjustment regarding six different covariance structures.

Variables	Criteria	Structures					
		AR	CS	VC	HAR	HCS	UN
FM	AIC	47052.8	48623.8	48864.8	*	*	*
	BIC	47064.8	48635.9	48753.9	*	*	*
NF	AIC	32715.9	35501.4	36321.5	*	*	*
	BIC	32727.9	35513.5	36319.4	*	*	*
TY	AIC	64176.7	66832.5	67502.2	*	*	*
	BIC	64188.7	66844.5	67485.9	*	*	*

FM = fruit mass; NF = number of fruits, and TY = total yield per plant. AR = First-order autoregressive; CS = compound symmetry; VC = variance components; HAR = heterogeneous autoregressive; HCS = heterogeneous compound symmetry; UN = unstructured. *Did not converge.

positive, in other words, a prerequisite. One way to deal with convergence problems could be to use nonlinear models, generalized mixed models, and the Bayesian methods. The SAS software can use a Fisher score linked to the estimate method until a predetermined iteration number (Littell et al., 2006). It was decided, however, to avoid convergence because the models that did not converge contained many parameters in the covariance matrix and such hyper-parameterization is not wanted.

The AIC and BIC enable comparing models with different factors and provide a higher grade to the more parameterized models, which assigns a lower model adjustment. Among these criteria, the BIC is the strictest, as it is a criterion that favors models with the least possible parameters to be estimated (Wolfinger, 1993). If the HAR, HCS, and UN models had converged, they would not be selected because of the significant number of parameters of the model to be estimated. Besides, the model selection criteria score better for simpler models, harming the more complex ones.

The unstructured covariance structure is the most complex one because it estimates unique correlations for each pair of points in time, a hyper-parameterized model. This structure is not commonly used, which makes it an unusual model for perennial crops. In contrast, the compound symmetry structure is also a very parameterized covariance matrix, but it is used in some works in which AR and CS are also cited (Maia et al., 2013; Quintal et al., 2017).

When a possible existence of a linear correlation between the measurements of the experimental unit is neglected, a more significant error in the residual variance component will possibly be attributed, as everything that is not in the model goes to the residue (Islam and Chowdhury, 2017). This is seen when the simple variance components are used, which consider a zero covariance between the measurements, not representing the relationship between the measurements (Woyann et al., 2018).

The variance component (VC) used in simple repeatability models presented the highest values of AIC and BIC, the least true, as predicted. This was likely to occur because this covariance structure assumes a lack of correlation between measurements. Hence, the assumption of independence cannot be admitted as a rule to support the classical variance analysis model for this study (Silva et al., 2021).

The autoregressive structure had the lowest value for all traits in both selection criteria. In this covariance structure, correlations among observations of the same individual diminish throughout time. In other words, correlations among the observations of the first harvest are greater with the second harvest, are smaller with the third, and are much smaller with the fourth harvest. Therefore, this structure was the most suitable to represent the existing correlation between the measurements according to the adjustment of the models. Working with three harvests, Quintal et al.

(2017) concluded that the most appropriate structures to model yield variables in guava tree crops were AR and CS respectively. Similar results have shown that the spatial modeling of errors in Pear orange clones can be done by using first-order autoregressive model. This covariance structure enabled a better fit among the models under evaluation (Maia et al., 2013).

The covariance matrix parameters were estimated ($\hat{\rho}$) for the structure with the lowest AIC and BIC values, the autoregressive. The values estimated were 0.57, 0.87, and 0.87 for the variables fruit mass, number of fruits, and total yield per plant, respectively.

The parameter ρ estimated for the AR covariance structure approaches zero as harvests pass. For example, when plotting the response of yield along the time, the primary variable of interest in the crop, it can be seen that the climatic conditions influenced each crop at that time. Hence, as time passes, although measurements are taken in the same location, the climate of a measurement does not have much influence on another measurement taken long after (Figure 1).

After selecting the best model, the repeatability coefficient was estimated (Table 2). This coefficient was estimated for both variables, fruit mass and total plant yield, which presented repeatability estimates of 0.24 and 0.54, respectively. This variation in repeatability coefficients may be related to the nature of the traits, the genetic properties of the population, and whether the individuals under evaluation are stabilized (Cruz et al., 2012).

The repeatability coefficient estimates are considered high when they are equal to or higher than 0.6; median when the estimates display values between 0.6 and 0.3; and low when the values are below 0.3 (Resende et al., 2006). Variables with a repeatability coefficient above 0.5 with a coefficient of determination above 80 % prove the reliability of the phenotypic value to predict the true value of individuals (Bergo et al., 2013).

The variables number of fruits and total yield per plant showed estimates above 0.5, considered moderate values. Similar values were found by Costa (2003) by testing different methods to estimate the repeatability coefficient and working with the same variables in mango. The authors concluded that the coefficients estimated suggested that the environmental variance had little influence on these variables from one harvest to the other.

Pruning is a phytosanitary treatment that greatly influences these variables. In guava trees, yield pruning and subsequent removal of sprouts are commonly made, keeping the amount of branching, which will produce floral buds. If many branches are kept during sprout thinning, the number of fruits increases, but the fruit mass diminishes due to the distribution of the plant resources. Hence, the "sprout thinning environment" performed by the breeder influences the variables, but it must be constant throughout time to avoid influencing parameters such as heritability.

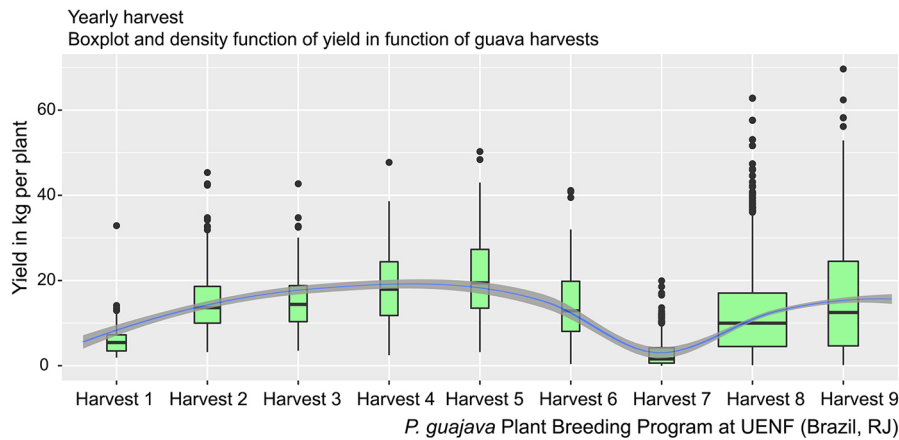


Figure 1 – Yield profile in *Psidium guajava* for nine harvests. The blue line is the density function of 17 guava tree families throughout time. The area around the blue line is the standard error. The green boxplot is the quantile and median of 17 guava tree families at a time.

Table 2 – Accuracy (A), efficiency (E), and repeatability (R²) values for the variables fruit mass (FM), number of fruits (NF), and total yield per plant (TY) in guava tree population.

Repeatability	FM			NF			TY		
	R ²	A	E	R ²	A	E	R ²	A	E
	0.24			0.54			0.53		
Harvests	R ²	A	E	R ²	A	E	R ²	A	E
1	0.24	0.49	1	0.54	0.74	1	0.53	0.73	1
2	0.39	0.62	1.26	0.70	0.84	1.13	0.69	0.83	1.14
3	0.49	0.70	1.41	0.78	0.88	1.19	0.77	0.88	1.20
4	0.56	0.75	1.51	0.83	0.91	1.22	0.82	0.90	1.23
5	0.61	0.78	1.58	0.85	0.92	1.25	0.85	0.92	1.26
6	0.66	0.81	1.64	0.87	0.93	1.26	0.87	0.93	1.27
7	0.69	0.83	1.68	0.89	0.94	1.27	0.88	0.94	1.28
8	0.72	0.84	1.71	0.90	0.95	1.28	0.90	0.94	1.29
9	0.74	0.86	1.74	0.91	0.95	1.29	0.91	0.95	1.30

Regarding the estimated number of measurements, four harvests can be established as the minimum number of observations required in the same individual for the variables number of fruits and total yield per plant. These measurements can lead to reliable data that enables individual selection, with over 90 % reliability and minimal cost and labor. Quintal et al. (2017) and Almeida et al. (2019) worked with three guava crops to estimate the number of harvests needed with predictions from the third harvest. The authors concluded that a more significant number of harvests would be necessary, five of them, to reach a sure accuracy. In our study, however, in which measurements were performed and not predicted, it was recommended four harvests.

The repeatability in terms of the mean level of the four harvests (coefficient of determination) corresponds to 0.83 (NF) and 0.82 (TY). Coefficients of determination greater than 0.8 prove the reliability of the phenotypic value in predicting the true value of this population. A repeatability study with mango yield variables reported similar estimates (Costa, 2003).

The individual accuracy was 0.90 for the variables NF and TY. The selective accuracy results from the estimate of heritability, repeatability of the variable, and methodologies to predict genetic values. Given that this measurement is linked to the correlation between predicted genetic values and true genetic values of individuals, the greater the accuracy in the evaluation of an individual, the greater the reliability in the evaluation of the individual.

The efficacy of five harvests compared to only one is 1.22 and 1.23 for the variables NF and TY, meaning that when four harvests were used, an increase of more than 20 % in efficacy was obtained on average compared to one. From the fourth harvest onward, the increase in the number of harvests presented a slight gain in efficacy; thus, an increase in the number of harvests was not viable.

The trait fruit mass showed an estimated repeatability of 0.24, indicating low regularity in the repetition of the character from one cycle to the other. In order to determine a minimum number of measurements to predict the true value of individuals, based on the

coefficient of determination of 0.85, considered a reliable value for a trait with low heritability (Bergo et al., 2013) for this variable, it was used:

$$n_0 = \frac{R^2(1-r)}{(1-R^2)r} = \frac{0.85(1-0.24)}{(1-0.85)0.24} = 17.94$$

Hence, to reach a coefficient of determination of 85 %, it is necessary to perform approximately 18 measurements per individual. Therefore, the application of breeding methods that have a good parental control of the individuals is required to obtain gains in this variable, as well as indirect selection by means of studying correlations with better genetic control traits (Maia et al., 2013).

This low repeatability coefficient for the variable FM may be attributed to the genetic difference between the genotypes analyzed in the experiment, the experimental control, and the environmental variations due to the long period of exposure of the plants to the environment (perenniality). Nevertheless, this variable is relevant for fruit tree breeding, despite being a trait of low repeatability.

In this case, indirect selection by studying correlations can be a good strategy. For example, pulp mass, fruit diameter, and fruit length exhibited a correlation of 0.95, 0.9, and 0.78 with fruit mass (Silva et al., 2021) and can be used to select this variable indirectly.

Given the results obtained for this variable, some points should be emphasized. This variable does not keep the means similar throughout the harvests, which results in a low repeatability value. This low value for repeatability influences the model, indicating the need to carry out 18 harvests to make the selection in it. When a genetic breeding program is conducted on perennial plants, it is impossible to perform this number of harvests. Additional works have already evidenced that it is not viable to increase the number of measurements to reach higher levels of accuracy in perennial plant variables, such as mango tree crops (Costa, 2003).

Therefore, yield, the main trait of interest in guava tree crops, can be evaluated with few harvests, which still allows obtaining reliable data for individual predictions. Other traits, such as fruit mass, commonly sought for cultivars intended for table fruit, can be selected using indirect selection by other correlated variables.

Conclusion

The autoregressive structure provided the best results for all the variables, thus being suitable for modeling this type of experiment.

Four measurements can be used to estimate a value close to the true value of individuals for the variables NF and TY.

The variable FM provided a low repeatability coefficient value, requiring further observations to obtain high accuracy, making it impossible to rely solely on this variable.

Authors' Contributions

Conceptualization: Silva, F.A.; Viana, A.P. **Data curation:** Silva, F.A.; Leal, L.S.S. **Formal analysis:** Silva, F.A.; Correa, C.C.G. **Methodology:** Glória, L.S. **Writing – review & editing:** Silva, F.A.; Viana, A.P.

References

- Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions On Automatic Control* 19: 716-723. <https://doi.org/10.1109/tac.1974.1100705>
- Almeida, C.L.P.; Viana, A.P.; Santos, E.A.; Quintal, S.S.R. 2019. Repeatability in guava: how many evaluations is necessary for selection the best guava tree? *Functional Plant Breeding Journal* 1: 5. <http://dx.doi.org/10.35418/2526-4117/v1n2a5>
- Bergo, C.L.; Negreiros, J.R.S.; Miqueloni, D.P.; Lunz, A.M.P. 2013. Repeatability estimates of yield traits in peach palm to palm heart of Putumayo landrace. *Revista Brasileira de Fruticultura* 35: 829-836 (in Portuguese, with abstract in English). <https://doi.org/10.1590/S0100-29452013000300020>
- Costa, J.G. 2003. Estimate of repeatability of some traits of production in mango tree. *Ciência Rural* 33: 263-266 (in Portuguese, with abstract in English). <http://dx.doi.org/10.1590/S0103-84782003000200013>
- Cruz, C.D.; Regazzi, A.J.; Carneiro, P.C.S. 2012. *Biometric Methods Applied to Genetic Improvement = Modelos Biométricos Aplicados ao Melhoramento Genético*. 2nd Edition, UFV, Viçosa. (in Portuguese)
- Islam, M.A.; Chowdhury, R.I. 2017. *Analysis of Repeated Measures Data*. Springer Nature: London, UK. <https://doi.org/10.1007/978-981-10-3794-8>
- Littell, R.C.; Milliken, G.A.; Stroup, W.W.; Wolfinger, R.D.; Oliver, S. 2006. *SAS for Mixed Models*. SAS, Cary, NC, USA.
- Maia, E.; Siqueira, D.L.; Carvalho, S.A.; Peternelli, L.A.; Latado, R.R. 2013. Application of the spatial analysis on the evaluation of selection experiments of Pera orange tree clones. *Ciência Rural* 43: 8-14. (in Portuguese, with abstract in English) <http://dx.doi.org/10.1590/S0103-84782012005000134>
- Mathew, B.; Leon, J.; Sillanpää, M.J. 2018. Impact of residual covariance structures on genomic prediction ability in multi-environment trials. *PloS One* 13: e0201181. <https://doi.org/10.1371/journal.pone.0201181>
- Patterson, H.D.; Thompson, R. 1971. Recovery of inter-block information when block sizes are unequal. *Biometrika* 58: 545-554. <https://doi.org/10.1093/biomet/58.3.545>
- Quintal, S.S.R.; Viana, A.P.; Campos, B.M.; Vivas, M.; Amaral-Júnior, A.T. 2017. Analysis of structures of covariance and repeatability in guava segregating population. *Revista Caatinga* 30: 885-891. <https://doi.org/10.1590/1983-21252017v30n408rc>
- Resende, M.D.V.; Thompson, R.; Welham, S. 2006. Multivariate spatial statistical analysis of longitudinal data in perennial crops. *Revista de Matemática e Estatística* 24: 147-169.

- Resende, R.T.; Resende, M.D.V.; Silva, F.F.; Takahashi, E.K. 2017. Predictive accuracy of *Eucalyptus* spp. clonal trials using additive kinship effects and cross validation. *Scientia Forestalis* 45: 39-47 (in Portuguese, with abstract in English). <https://dx.doi.org/10.18671/scifor.v45n113.04>
- Schwarz, G. 1978. Estimating the dimension of a model. *The Annals of Statistics* 6: 461-464. <https://doi.org/10.1214/aos/1176344136>
- Shalizi, M.N.; Isik, F. 2019. Genetic parameter estimates and GxE interaction in a large cloned population of *Pinus taeda* L. *Tree Genetics & Genomes* 15: 46. <https://doi.org/10.1007/s11295-019-1352-7>
- Silva, F.A.; Correa, C.C.G.; Carvalho, B.M.; Viana, A.P.; Preisigke, S.C.; Amaral-Júnior, A.T. 2021. Novel approach to the selection of *Psidium guajava* genotypes using latent traits to bypass multicollinearity. *Scientia Agricola* 78: e20190081. <https://doi.org/10.1590/1678-992x-2019-0081>
- Wade, K.M.; Quaas, R.L. 1993. Solutions to a system of equations involving a first-order autoregressive process. *Journal of Dairy Science* 76: 3026-3032. [https://doi.org/10.3168/jds.s0022-0302\(93\)77642-0](https://doi.org/10.3168/jds.s0022-0302(93)77642-0)
- Wolfinger, R. 1993. Covariance structure selection in general mixed models. *Communications in Statistics-Simulation and Computation* 22: 1079-1106. <https://doi.org/10.1080/03610919308813143>
- Woyann, L.G.; Milioli, A.S.; Bozi, A.H.; Dalló, S.C.; Matei, G.; Storck, L.; Benin, G. 2018. Repeatability of associations between analytical methods of adaptability, stability, and productivity in soybean. *Pesquisa Agropecuária Brasileira* 53: 63-73. <https://doi.org/10.1590/s0100-204x2018000100007>