



Interrelations between productive, vegetative, and fruit quality traits in Tahiti acid lime through Bayesian inference

Renan Garcia Malikouski¹  Emanuel Ferrari do Nascimento¹  Andréia Lopes de Morais² 
Marco Antônio Peixoto¹  Moises Zucoloto³  Dimmy Herllen Silveira Gomes Barbosa⁴ 
Leonardo Lopes Bhering^{1*} 

¹Departamento de Biologia Geral, Universidade Federal de Viçosa (UFV), 36570-900, Viçosa, MG, Brasil. E-mail: leonardo.bhering@ufv.br.
*Corresponding author.

²Departamento de Ciências Agrárias e Biológicas, Universidade Federal do Espírito Santo (UFES), São Mateus, ES, Brasil.

³Departamento de Agronomia, Universidade Federal do Espírito Santo, Alegre, ES, Brasil

⁴Embrapa Mandioca e Fruticultura Tropical, Cruz das Almas, BA, Brasil.

ABSTRACT: *Although the fruit yield has a core importance in Tahiti acid lime breeding programs, other traits stand out among the quality fruit and vegetative traits as ones that still need to be improved in selection of superior genotypes. Applying efficient tools aiming selection, such as the Bayesian inference, becomes an alternative in perennial crops. This study applied Bayesian inference in the genetic evaluation of Tahiti acid lime genotypes and estimated the interrelation between vegetative, productive and fruit quality traits. Twenty-four acid lime genotypes were evaluated for number of fruits, fruit yield, canopy volume, stem diameter, soluble solids content, shell thickness, and juice yield traits. The genotypic values were estimated through Bayesian inference and models with different residual structure were tested via deviance information criterion. Pearson's correlation and the path analysis were estimated, removing the multicollinearity effect. The Bayesian inference estimates genotypic values with high selective accuracy. The correlations obtained between traits from different groups can be useful in selection strategies for improvement of Tahiti acid lime. The Bayesian inference demonstrated to be an important tool and should be considered in perennial breeding programs.*

Key words: *Citrus latifolia Tanaka, longitudinal data, Markov chain Monte Carlo, Pearson's correlation, direct and indirect effects.*

Inter-relações entre características produtivas, vegetativas e de qualidade de fruto em lima ácida Tahiti através de inferência Bayesiana

RESUMO: *Embora a produtividade seja uma característica fundamental em programas de melhoramento da lima ácida Tahiti, outras características se destacam por proporcionar uma seleção mais eficiente ou mesmo influenciar na expressão de atributos produtivos. A aplicação de inferência Bayesiana, pode tornar o processo de identificação de indivíduos superiores e o estudo das correlações entre as características ainda mais acurado. Objetivou-se, por meio deste estudo, estimar valores genéticos em genótipos de lima ácida Tahiti através de inferência Bayesiana e estimar coeficientes de correlação e de trilha entre características vegetativas, produtivas e de qualidade de frutos. Vinte e quatro genótipos de lima ácida foram avaliados durante dois anos para peso e número de frutos, volume de copa, diâmetro de caule, teor de sólidos solúveis dos frutos, espessura de casca e rendimento de suco. Os valores genéticos foram estimados por meio de inferência Bayesiana e modelos com diferentes estruturas residuais foram testados via critério de informação de deviance. Coeficiente de correlação de Pearson e de análise de trilha foram determinados, removendo o efeito de multicolinearidade. A estimação de valores genéticos via inferência Bayesiana apresentou alta acurácia seletiva. As correlações obtidas entre caracteres de diferentes grupos podem ser úteis em estratégias de seleção para melhoramento da lima ácida Tahiti. A inferência Bayesiana demonstrou ser uma ferramenta importante e deve ser considerada em programas de melhoramento de culturas perenes.*

Palavras-chave: *Citrus latifolia Tanaka, dados longitudinais, cadeias de Markov Monte Carlo, correlação de Pearson, efeitos diretos e indiretos.*

INTRODUCTION

Belonging to the *Rutaceae* family, the genus *Citrus* is one of the most important fruit crops in the world agricultural scenario (LADANIYA et al., 2020). Characterized by their high nutritional

value, *Citrus* fruits are consumed fresh or used in the preparation of food and drinks, while their essential oils are used in the manufacture of perfumes, cosmetics, and even medicines (DENARO et al., 2020; DOSOKY & SETZER, 2018). Among all species from *Citrus*, Tahiti acid lime (*Citrus latifolia*

Tanaka L.) has stood out in Brazil due to its good adaptation to edaphoclimatic conditions and high economic value (BASTOS et al., 2017). In view of the growing internal and external demand for fresh fruits in the citrus market, especially for Tahiti acid lime, it is essential to maximize the productivity of orchards (MORAIS et al., 2020).

The selection of superior genotypes in terms of fruit yield is an important way to contribute to increased production in *Citrus* species (IMAI et al., 2016). However, fruit yield is a quantitative and complex trait aiming for selection since it has low heritability, and it is a result of the association of several productive components (LI et al., 2017). The knowledge of the correlations between pairs of traits allows to identify traits suitable to conduct indirect selection, resembling great importance when the direct selection of target trait is hampered by low heritability or high measurement cost (CREVELARI et al., 2018). However, the correlation coefficients alone do not provide the direct and indirect effects under fruit yield, being path analysis an excellent strategy for such end. It establishes cause and effect information between correlations and has been applied to different fruit crops of economic importance such as cocoa (DEVY et al., 2018), guajava (SANTOS et al., 2017), peach (MATIAS et al., 2016), and grape (DOLKAR et al., 2017).

The mixed model methodology is considered the standard methodology in perennial plants (OLIVEIRA et al., 2014; SILVA et al., 2020). However, it has some statistical limitations, such as providing only approximated confidence intervals and the assumption of asymptotic normality (RESENDE, 2000). Thus, the Bayesian inference overcomes such problems and presents a great alternative in relation to frequentist analysis, as it allows the use of a priori information about the model parameters and the estimation of exactly credibility intervals (RESENDE, 2000).

Although its potential, for Tahiti acidic lime there is no study involving estimation of genotypic values through Bayesian inference and its application in genetic selection, as well as the use of the values for estimation of correlations and coefficients of path analysis between traits. In view of the above, the aims of this study were to: i) implement the Bayesian inference for genetic assessment of Tahiti acid lime; ii) obtain estimates of correlation and direct and indirect effects among the main traits in Tahiti Lime based on Bayesian predicted genotypic values and; iii) Explore the estimates obtained through selection strategies in plant breeding.

MATERIALS AND METHODS

Experimental network and traits measurements

The experiment was implemented at São Mateus municipality, Espírito Santo State, Brazil (18°48'21"S, 39°53'30" W, 35 meters a. s. l.) and conducted between July 2017 and January 2019. Twenty-four Tahiti acid lime genotypes were grown at a spacing of 3 m between plants and 6 m between lines, using a randomized complete block design to correct the heterogeneity of the area, with 4 replications. The experimental unit consisted of three plants per plot, totaling 288 individuals. Seven traits were measured divided into three groups: two traits from vegetative development group: stem diameter (SD, in mm), measured using a digital caliper five centimeters above the graft region; and canopy volume (CV, in m³) obtained through the equation proposed by ZEKRI (2000) $CV = \left(\frac{\pi}{6}\right) \times PH \times DR \times DS$ where *PH* is plant height, *DR* and *DS* are diameter of canopy projection on the line and between the cultivation lines respectively, *PH*, *DR* and *DS* where measured through a measuring tape in meters. The other group was productive group, with two traits measured: fruit yield per plant (FY, in kg plant⁻¹), where the fruits of each plant with a minimum diameter of 47 mm, smooth shell and light green color were harvested and weighed with the aid of a digital scale; and number of fruits per plant (NF) through counting. The third group was of fruit quality with three traits: shell thickness (ST, in mm) where the fruits were cut in half and the peel measured with the aid of a digital caliper; juice yield (JY, in %) calculated from the fruit weight/fruit bagasse weight ratio, and soluble solids content (SS, in °brix) verified with the portable digital refractometer type Palette PR - 32, brand ATAGO, with automatic temperature compensation. Four evaluations were carried out for the vegetative development group (at 736, 918, 1083 and 1249 days after planting), eight evaluations for productive group (at 736, 808, 861, 918, 972, 1083, 1200 and 1249 days after planting) and three evaluations in the fruit quality group (at 972, 1083 and 1200 days after planting).

Statistical analyses

The variance components and genotypic values were estimated by the Markov Chain Monte Carlo (MCMC) algorithm. The repeatability model, based on plot means, was determined according to QUINTAL et al. (2017), following the structure: $y = X\beta + Zg + W_i + T_s + e$, where *y* is the vector of phenotypic data, with the conditional distribution assumed as: $y | \beta, g, i, s, \hat{\sigma}_g^2$

$\hat{\sigma}_i^2$, $\hat{\sigma}_s^2$, $R \sim N(X\beta + Zg + Wi + Ts + e, R \otimes I)$, where $\hat{\sigma}_g^2$ is the genotypic (co)variance matrix, $\hat{\sigma}_i^2$ is the genotype by measurement ($G \times M$) interaction, $\hat{\sigma}_s^2$ is the permanent environmental (co)variance matrix, R is the residual (co)variance matrix, and I is an identity matrix; β is the vector of measurement-replication combinations (systematic effect), which comprises the effects of measurement, replicates and replication \times measurement interaction, assumed as $\beta | \mu_\beta, 1_\beta \hat{\sigma}_\beta^2 \sim N(\mu_\beta, 1_\beta \hat{\sigma}_\beta^2)$; g is the vector of genotypic effects, assumed as $g | \hat{\sigma}_g^2 \sim N(0, \hat{\sigma}_g^2 \otimes I)$; i is the vector of $G \times M$ effects, assumed as $i | \hat{\sigma}_i^2 \sim N(0, \hat{\sigma}_i^2 \otimes I)$; s is the vector of permanent environmental effect, assumed as $s | \hat{\sigma}_s^2 \sim N(0, \hat{\sigma}_s^2 \otimes I)$. The terms X , Z , W , and T refer to the incidence matrix for the systematic, genotypic, $G \times M$ interaction, and permanent environmental effects, respectively.

We assumed that $\hat{\sigma}_g^2$, $\hat{\sigma}_i^2$, $\hat{\sigma}_s^2$, and R follow an inverted Wishart distribution $WI(v, V)$, with hyperparameters v and V (SORENSEN & GIANOLA, 2002). The hyperparameters for all prior distributions were selected to provide non-informative or flat prior distribution. For the systematic effects β , a uniform prior distribution was attributed.

Furthermore, the parameters β , g , i , s , $\hat{\sigma}_g^2$, $\hat{\sigma}_i^2$, $\hat{\sigma}_s^2$, and R were estimated, following the joint posterior distribution:

$$P(\beta, g, i, s, \hat{\sigma}_g^2, \hat{\sigma}_i^2, \hat{\sigma}_s^2, R | y) \propto P(y | \beta, g, i, s, \hat{\sigma}_g^2, \hat{\sigma}_i^2, \hat{\sigma}_s^2, R) \times P(\beta, g, i, s, \hat{\sigma}_g^2, \hat{\sigma}_i^2, \hat{\sigma}_s^2, R).$$

Modelling residual variance

Models accounting for different residual variance structure (diagonal – DIAG and variance component – IDV) were tested through deviance information criterion (DIC) (SPIEGELHALTER et al., 2002) for each trait. The DIC was calculated as follows: $DIC = D(\theta) + 2pD$, where $D(\theta)$ is a point estimate of the deviance obtained by replacing the parameters with their posterior mean estimates in the likelihood function, and pD is the effective number of estimated parameters. Also, DIC was used to test the significance of random effects of the model.

The MCMC algorithm through Gibbs sampling, was implemented as an iterative process. For the iterative estimation process, a total of 2.000.000 iterations were considered, being 200.000 chains considered as the burn-in period and an interval of sampling from 10 (thin period) were used. As a total, 180.000 chains were performed. After concluding the iterative process, the chain's convergence was checked by HEIDELBERGER & WELCH (1983) criterion. Further, the mean selective

accuracy (\tilde{r}_{soi}) was estimated as (RESENDE et al., 2014): $\tilde{r}_{\text{soi}} = 1 - \frac{s(\hat{g}_i)}{\hat{g}_i}$ where $s(\hat{g}_i)$ stands for the standard deviation of the estimated genotypic value.

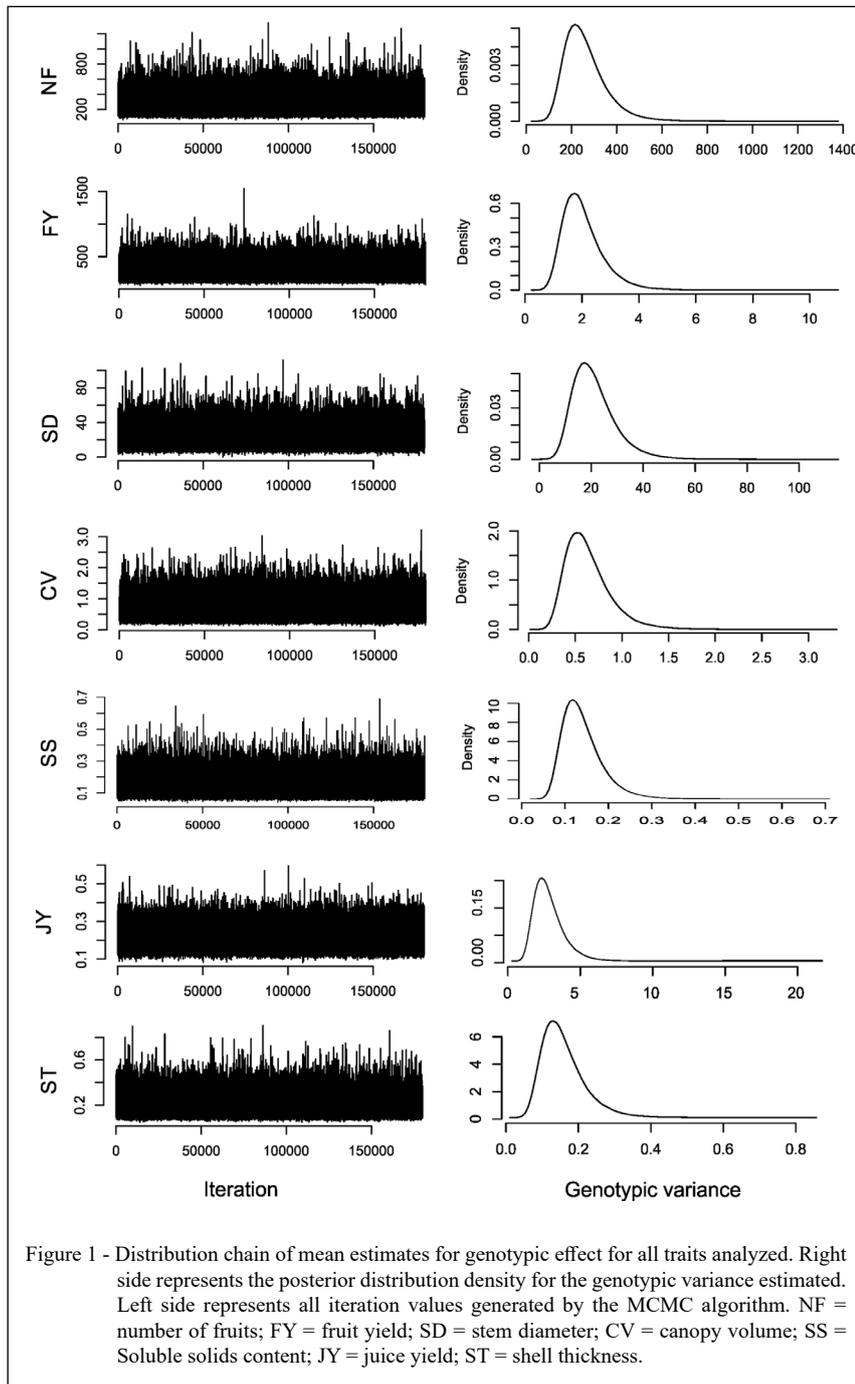
Path analysis and correlations

The genotypic values estimated using the means of MCMC algorithm were; therefore, subjected to Pearson's correlation analysis, with the verification of the coefficient's significance via the t test, with a 5% error probability. The path analysis was preceded by a diagnosis of multicollinearity test. For such an end, the estimates of correlation coefficients between the explanatory variables through the analysis of the condition number (CN). The path analysis was performed in two different scenarios. The first scenario was carried out considering the FY trait as the main trait due to its economic importance and the polygenic nature, whereas the second scenario was considering the SS trait as the main one, because it is a trait that needs specific equipment for measuring. The Bayesian analyses were carried out in the R software (R DEVELOPMENT CORE TEAM, 2020) through the packages 'MCMCglmm' (HADFIELD, 2010) and 'boa' (SMITH, 2007). The Pearson's correlations and the path analysis were carried out through the Genes software (CRUZ, 2016).

RESULTS AND DISCUSSION

The diagnostic of chain convergence (the Heidelberg stationarity test), indicated that all (co)variance components converged when 180,000 effective samples were used (Figure 1). In addition, the values of DIC indicated the model with residual variance as diagonal the best fit for all traits analyzed (Table 1). In perennial breeding better fitted models present, in general, heterogeneous residual variances structures, such as diagonal or unstructured (GIRI et al., 2019). For instance, spatial variation is recognized as the main disturbing factor for perennial crops and it may differ as a result of seasonal changes and crop growth stage (FAVERI et al., 2015; RESENDE, 2002). Results evidenced that all effects in the model were significant by DIC, except for the traits from the fruit quality group (SS and ST traits) (Table 1). Then, the Bayesian model was able to predict the genotypic values from the traits analyzed (Figure 2). The high posterior density interval (HPD) obtained by Bayesian inference tends to be shorter or leastwise equal than confidence interval, and should be preferred (SILVA et al., 2020).

The selective accuracy values ranging from 0.69 to 0.95 (NF: 0.83; FY: 0.84; SD: 0.69; CV:



0.92; SS: 0.94; JY: 0.95; and, ST: 0.91). According to RESENDE & ALVES (2020) the selective accuracy values were considered adequate, allowing selection of genotypes with superior performance and interrelation studies among traits.

Higher correlations values were observed between the pair or traits NF and FY (0.99). In

addition, both traits presented high correlation with JY trait (0.73 and 0.70 for NF and FY, respectively). The traits referring to the fruit quality group presented high correlation with each other (0.77 for JY x SS and 0.80 for ST x SS), in exception to the pair of traits JY x ST, where the correlation was 0.45. All traits from the productivity group (NF and FY) showed a non-

Table 1 - Deviance inference criterion for a homogeneous (IDV) and heterogeneous (DIAG) model of residual variance and significance of effects for the chosen model for all traits analyzed. NF = number of fruits; FY = fruit yield; SD = stem diameter; CV = canopy volume; SS = Soluble solids content; JY = juice yield; ST = shell thickness.

Structure	Trait						
	NF	FY	SD	CV	SS	JY	ST
IDV	7283.92	3567.27	2747.23	1205.43	450.15	1817.73	227.338
DIAG	7050.25 [#]	3258.72 [#]	2610.58 [#]	1015.87 [#]	434.91 [#]	1808.98 [#]	209.462 [#]
Effect	Effects significance test						
Genotypic	7078.69	3296.22	2612.38	1026.37	435.576	1807.47*	209.88*
Interaction	7086.08	3384.36	2619.4	1094.95	581.05	1878.83	358.07
Permanent	7173.95	3316.85	2676.91	1149.63	438.052	1829.09	213.11
Full model	7050.25*	3258.72*	2610.58*	1015.87*	434.911*	1808.98	209.46

[#] indicates the residual error variance that presented the best fit to the data.

* model with the best adjust based on DIC.

zero correlation with the traits from the fruit quality group (NF x SS: 0.45; NF x FY: 0.99, and SS x FY: 0.46). However, none of the traits presented in these two groups showed any significant correlation with the traits from the vegetative group (CV and SD), whereas the correlation between these two was the second highest reported in the analyses (0.86). The other correlations were not statistically different from zero.

These results highlight the higher correlations between pairs of traits from the same group. It is a result of the same genes controlling those traits, which was not indicated for pairs of traits from different groups. The exception, JY trait, presented high correlation with the traits from the productive group. However, this trait can be considered as a direct response of the other ones (FY and NF), implying in the high correlation observed.

Multicollinearity diagnosis identified severe multicollinearity for the matrix of the genotypic values estimated through Bayesian inference (NC > 100) for both scenarios (FY and SS as the main trait). When the trait NF was taken out from the analysis, multicollinearity diagnosis analysis obtained was considered weak in both cases (NC equal to 34.04 and 27.18 for FY and SS as the main trait, respectively) (Table 2).

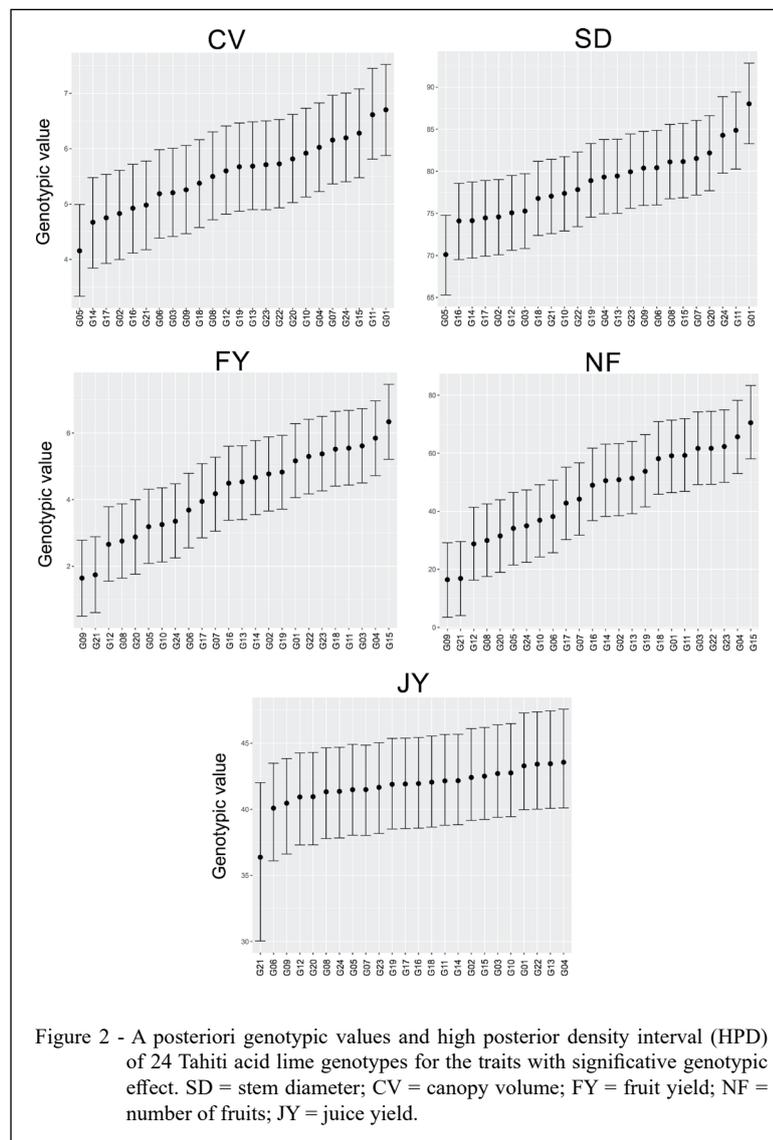
The path analysis considering FY as the main trait revealed a positive direct effect for CV (0.2306) and JY (0.7740). Conversely, traits SD, ST, and SS showed negative and lower values of direct effect (-0.1392, -0.0333 and -0.1474, respectively) (Table 2). As for the indirect effects, the contribution of SS via JY (0.6009) stands out. In the total sum of the effects (direct and indirect) there was a greater contribution of the JY trait (0.7065) and a lower

contribution of SD trait (0.1142) in the expression of FY. The quality parameters of the analysis showed a moderate coefficient of determination (R^2) of 53.45% and a residual effect of 0.6823 (Table 2).

Similar results were observed in the second scenario (SS as the main trait). The highest value of positive direct effect was reported by JY1 trait (0.5913). The ST trait demonstrated a direct effect of similar magnitude (0.5409). For indirect effects in the expression of SS trait, the greatest contribution was observed for SD via JY (0.45), JY through the SD trait (0.4178). In addition, they were higher than the residual effect (0.37), whereas the coefficient of determination was considered higher (85.99%) (Table 2).

The JY trait was responsible for the greatest direct effect in both scenarios. In addition, it presented significant genotypic correlation with FY and SS, showing that there is a good occurrence of cause and effect relationship. It can be said that JY is important in explaining FY and SS in addition to being a relevant feature in determining the quality of the fruit (DUBEY & SHARMA, 2016). Greater efficiency was observed in the path analysis considering SS as the main one, since FY, SD, CV, ST and JY explained 85.99% of SS, in addition to presenting a residual effect lower than the scenario that considered FY as the main trait. This trait requires specialized equipment to perform the measurement; a good explanation of SS based on the direct and indirect effects of the secondary traits may allow its exclusion in evaluations and future research, which provides savings in resources, infrastructure and labor.

In exception of JY in both scenarios and ST in the scenario in which SS was considered the main



trait, there were low direct effects on the other traits and a high genotypic correlation between the traits of production and fruit quality. This is an indication that the truncated selection in the auxiliary trait may not provide satisfactory gain with the selection, being the simultaneous trait selection strategy a better one to be employed (CRUZ et al., 2014). Despite not having a significant genotypic effect, JY shows to be a great alternative for indirect selection in genotypes that present genotypic variation for the referred trait.

Although FY trait is an easy trait to measure, the path analysis accounting for the FY as main trait demonstrates not to be suitable for this study, once it provides a lower determination coefficient of the

regression and higher residual effect. It differs from the SS trait as the main scenario. The discussion regarding the best path analysis for the study of Tahiti acid lime was validated by these two indicators. Furthermore, the choice of SS trait is the most accepted.

These results show the importance of carrying out analyzes of the relationship between traits, since the productive performance of crops involves other factors that can influence more or less incident in their expression. In addition, to measure many traits require sophisticated equipment for their measurement, and the identification of easier and cheaper alternatives for characterizing genotypes can considerably decrease the costs of a breeding program.

Table 2 - Path coefficient analysis showing direct and indirect effects on the fruit yield (FY) (Scenario 1) and soluble solids content (SS) (Scenario 2) traits.

Effects	Traits ⁽¹⁾					
	Scenario 1					
	SD	CV	ST	SS	JY	
Direct on FY	-0.1392	0.2306	-0.0333	-0.1474	0.7740	
Indirect via SD	--	-0.1209	-0.0036	-0.0046	-0.0106	
Indirect via CV	0.2003	--	0.0038	0.0309	0.0729	
Indirect via ST	-0.0009	-0.0005	--	-0.0267	-0.0152	
Indirect via SS	-0.0048	-0.0197	-0.1183	--	-0.1145	
Indirect via JY	0.0589	0.2445	0.3537	0.6009	--	
Total	0.1142	0.3339	0.2023	0.4530	0.7065	
Effects	Scenario 2					
	FY	SD	CV	ST	JY	
Direct on SS	-0.0444	0.0802	-0.1167	0.5409	0.5913	
Indirect via FY	--	-0.0051	-0.0148	-0.0090	-0.0314	
Indirect via SD	0.0092	--	0.0697	0.0021	0.0061	
Indirect via CV	-0.0390	-0.1013	--	-0.0019	-0.0369	
Indirect via ST	0.1094	0.0141	0.0090	--	0.2472	
Indirect via JY	0.4178	0.4500	0.1868	0.2702	--	
Total	0.4530	0.0329	0.1339	0.8022	0.7763	
R ² (%) ⁽²⁾	53.45			85.99		
Residual Effect ⁽²⁾	-0.6823			-0.3743		
NC ⁽³⁾	34.04			27.18		

⁽¹⁾Stem diameter (ST), canopy volume (CV), shell thickness (ST) and juice yield (JY) on soluble solids content (SS).

⁽²⁾Determination coefficient (R²) and Residual effect for scenario 1 (left) and 2 (right), respectively.

⁽³⁾Multicollinearity diagnosis. Number of condition (NC) obtained by the ratio between the highest and smallest eigenvalue of the correlation matrix for scenario 1 (left) and 2 (right), respectively. Value less than 100 is not a problem for the analysis (MONTGOMERY; et al, 2012).

CONCLUSIONS

The values of selective accuracy demonstrated the reliability of Bayesian inference to estimation of genotypic values. The use of Bayesian approaches combined with correlation analysis and path analysis are effective strategies and should be considered in the statistical evaluation of breeding programs for Tahiti acid lime and other fruit and perennial species.

According to the path analysis, the JY trait present the large influence in the expression of FY, and the traits JY with ST influences in the expression of SS trait. The interrelation between these traits is important to determining which selection strategies are suitable to be applied in a breeding program.

DECLARATION OF CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

The authors are grateful to the Bello Fruit company, which provided the facilities for the development of the experiment, Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) and Frucafé company, the source of the genotypes. We are thankful to the Universidade Federal do Espírito Santo and to the Universidade Federal de Viçosa for their contribution for human resource training. We appreciate the financial support from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), the Espírito Santo Research and Innovation Support Foundation (FAPES), and the Fundação de Amparo a Pesquisa do Estado de Minas Gerais (Fapemig). This study was financed in part by the CAPES - Finance Code 001.

AUTHORS' CONTRIBUTIONS

All authors contributed equally for the conception and writing of the manuscript. All authors critically revised the manuscript and approved of the final version.

REFERENCES

BASTOS, D. C. et al. Initial development of Tahiti acid lime trees on different rootstocks in the semiarid region of Ceará,

- Brazil. **Citrus Research & Technology**, v.38, n.1, p.77–82, 2017. Available from: <<https://citrusrtr.ccsm.br/article/10.4322/crt.ICC045/pdf/citrusrtr-38-1-77.pdf>>. Accessed: Nov. 12, 2020. doi: 10.4322/crt.icc045.
- CREVELARI, J. A. et al. Correlations between agronomic traits and path analysis for silage production in maize hybrids. **Bragantia**, v.77, n.2, p.243–252, 2018. Available from: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S000687052018000200243&tln=en>. Accessed: Nov. 03, 2020. doi: 10.1590/1678-4499.2016512.
- CRUZ, C. D. Genes Software-extended and integrated with the R, Matlab and Selegen. **Acta Scientiarum. Agronomy**, v.38, n.4, p.547–552, 2016. Available from: <<https://doi.org/10.4025/actasciagron.v38i4.32629>>. Accessed: Oct. 15, 2020. doi: 10.4025/actasciagron.v38i4.32629.
- CRUZ, C. D. et al. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa: Editora UFV, 2014. 668p.
- DENARO, M. et al. New insights into Citrus genus: From ancient fruits to new hybrids. **Food Frontiers**, v.1, n.3, p.305–328, 2020. Available from: <<https://onlinelibrary.wiley.com/doi/epdf/10.1002/fft2.38>>. Accessed: Dec. 04, 2020. doi: 10.1002/fft2.38.
- DEVY, L. et al. Identification of Molecular Marker Based on MYB Transcription Factor for the Selection of Indonesian Fine Cacao (Theobroma cacao L.). **Pelita Perkebunan (a Coffee and Cocoa Research Journal)**, v.34, n.2, p.59–68, 2018. Available from: <<https://doi.org/10.22302/icri.jur.pelitaperkebunan.v34i2.314>>. Accessed: Oct. 19, 2020. doi: 10.22302/icri.jur.pelitaperkebunan.v34i2.314.
- DOLKAR, T. et al. Genetic variability and correlation studies in grapes (Vitis vinifera L.) in Leh District of Jammu and Kashmir. **Advances in Horticultural Science**, v.31, n.4, p.241–247, 2017. Available from: <<https://doi.org/10.13128/ahs-22376>>. Accessed: Oct. 22, 2020. doi: 10.13128/ahs-22376.
- DOSOKY, N. S.; SETZER, W. N. Biological activities and safety of *Citrus* spp. Essential oils. **International Journal of Molecular Sciences**, v.19, n.7, p.1–25, 2018. Available from: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6073409/>>. Accessed: Nov. 04, 2020. doi: 10.3390/ijms19071966.
- DUBEY, A.; SHARMA, R. M. Effect of rootstocks on tree growth, yield, quality and leaf mineral composition of lemon (Citrus limon (L.) Burm.). **Scientia Horticulturae**, v.200, p.131–136, 2016. Available from: <<https://doi.org/10.1016/j.scienta.2016.01.013>>. Accessed: Oct. 27, 2020. doi: 10.1016/j.scienta.2016.01.013.
- FAVERI, J. et al. Statistical methods for analysis of multi-harvest data from perennial pasture variety selection trials. **Crop and Pasture Science**, v.66, n.9, p.947–962, 2015. Available from: <<https://www.publish.csiro.au/cp/cp14312>>. Accessed: Nov. 09, 2020. doi: 10.1071/CP14312.
- GIRI, K. et al. Modelling and prediction of dry matter yield of perennial ryegrass cultivars sown in multi-environment multi-harvest trials in south-eastern Australia. **Field Crops Research**, v.243, p.107614, 2019. Available from: <<https://linkinghub.elsevier.com/retrieve/pii/S0378429019306471>>. Accessed: Nov. 05, 2020. doi: 10.1016/j.fcr.2019.107614.
- HADFIELD, J. D. MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. **Journal of Statistical Software**, v.33, n.2, p.1–22, 2010. Available from: <<http://www.jstatsoft.org/v33/i02/>>. Accessed: Nov. 10, 2020. doi: 10.18637/jss.v033.i02.
- HEIDELBERGER, P.; WELCH, P. D. Simulation run length control in the presence of an initial transient. **Operations Research**, v.31, p.1109–1144, 1983. Available from: <<https://doi.org/10.1287/opre.31.6.1109>>. Accessed: Dec. 01, 2020. doi: 10.1287/opre.31.6.1109.
- IMAI, A. et al. Evaluation of the best linear unbiased prediction method for breeding values of fruit-quality traits in citrus. **Tree Genetics & Genomes**, v.12, n.s6, p.119, 2016. Available from: <<https://link.springer.com/article/10.1007/s1129501610788>>. Accessed: Nov. 12, 2020. doi: 10.1007/s11295-016-1078-8.
- LADANIYA, M. S. et al. High density planting studies in acid lime (*Citrus aurantifolia* Swingle). **Scientia Horticulturae**, v.261, p.108935, 2020. Available from: <<https://linkinghub.elsevier.com/retrieve/pii/S0304423819308210>>. Accessed: Dec. 03, 2020. doi: 10.1016/j.scienta.2019.108935.
- LI, Y. et al. Genotype by environment interactions in forest tree breeding: review of methodology and perspectives on research and application. **Tree Genetics & Genomes**, v.13, n.3, p.60, 2017. Available from: <<http://link.springer.com/10.1007/s11295-017-1144-x>>. Accessed: Nov. 09, 2020. doi: 10.1007/s11295-017-1144-x.
- MATIAS, R. G. P. et al. Genetic diversity in peach cultivars. **Comunicata Scientiae**, v.7, n.3, p.293, 2016. Available from: <<https://www.comunicatascientiae.com.br/comunicata/article/view/1254/410>>. Accessed: Oct. 26, 2020. doi: 10.14295/cs.v7i3.1254.
- MONTGOMERY, D. C. et al. **Introduction to linear regression analysis**. New York: John Wiley and Sons, 2012. 672p.
- MORAIS, A. L. DE et al. Vegetative development and production of ‘Tahiti’ acid lime clone selections grafted on different rootstocks. **Revista Brasileira de Fruticultura**, v.42, n.3, e-585, 2020. Available from: <<https://doi.org/10.1590/0100-29452020585>>. Accessed: Oct. 28, 2020. doi: 10.1590/0100-29452020585.
- OLIVEIRA, E. J. et al. Genetic parameters and prediction of genotypic values for root quality traits in cassava using REML/BLUP. **Genetics and Molecular Research**, v.13, n.3, p.6683–6700, 2014. Available from: <<https://www.geneticsmr.com/articles/3476>>. Accessed: Nov. 05, 2020. doi: 10.4238/2014.August.28.13.
- QUINTAL, S. S. et al. Analysis of structures of covariance and repeatability in guava segregating population. **Revista Caatinga**, v.30, n.4, p.885–891, 2017. Available from: <<https://periodicos.ufersa.edu.br/index.php/caatinga/article/view/6211/pdf>>. Accessed: Apr. 10, 2021. doi: 10.1590/0100-29452018462.
- R DEVELOPMENT CORE TEAM. **R: A language and environment for statistical computing**. 2020. Online. Available from: <<https://www.r-project.org/>>. Accessed: Nov. 17, 2020.
- RESENDE, M. D. V. **Genética biométrica e estatística no melhoramento de plantas perenes**. Brasília: Embrapa Informação Tecnológica; Colombo: Embrapa Florestas, 2002. 975p.
- RESENDE, M. D. V. et al. **Estatística matemática, biométrica e computacional: modelos mistos, multivariados, categóricos e**

generalizados (REML/BLUP), inferência bayesiana, regressão aleatória, seleção genômica, QTL-QWAS, estatística espacial e temporal, competição, sobrevivência. Viçosa: UFV, 2014. 881p.

RESENDE, M. D. V. **Inferência bayesiana e simulação estocástica na estimação de componentes de variância e valores genéticos em plantas perenes**. Colombo: Embrapa Florestas, 2000. 68p.

RESENDE, M. D. V.; ALVES R. S. Linear, generalized, hierarchical, bayesian and random regression mixed models in genetics/genomics in plant breeding. **Functional Plant Breeding Journal**, v.2, p.1–31, 2020. Available from: <<https://doi.org/10.35418/2526-4117/v2n2a1>>. Accessed: Apr. 12, 2021. doi: 10.35418/2526-4117/v2n2a1.

SANTOS, W. N. L. Dos et al. Simultaneous determination of 13 phenolic bioactive compounds in guava (*Psidium guajava* L.) by HPLC-PAD with evaluation using PCA and Neural Network Analysis (NNA). **Microchemical Journal**, v.133, p.583–592, 2017. Available from: <<https://doi.org/10.1016/j.microc.2017.04.029>>. Accessed: Oct. 23, 2020. doi: 10.1016/j.microc.2017.04.029.

SILVA, F. A. DA et al. Impact of Bayesian Inference on the Selection of *Psidium guajava*. **Scientific Reports**, v.10, n.1, p.1999, 2020. Available from: <<http://www.nature.com/articles/>

s41598-020-58850-6>. Accessed: Oct. 18, 2020. doi: 10.1038/s41598-0205885-06.

SMITH, B. J. boa : An R Package for MCMC Output Convergence Assessment and Posterior Inference. **Journal of Statistical Software**, v.21, n.11, p.1–37, 2007. Available from: <<http://www.jstatsoft.org/v21/i11/>>. Accessed: Oct. 30, 2020. doi: 10.18637/jss.v021.i11.

SORENSEN, D.; GIANOLA, D. **Likelihood, Bayesian, and MCMC Methods in Quantitative Genetics**. New York: Springer-Verlag, 2002. 740p.

SPIEGELHALTER, D. J. et al. Bayesian measures of model complexity and fit. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, v.64, n.4, p.583–639, 2002. Available from: <<https://rss.onlinelibrary.wiley.com/doi/pdfdirect/10.1111/-1467-9868.00353>>. Accessed: Nov. 19, 2020. doi: 10.1111/1467-9868.00353.

ZEKRI, M. Citrus rootstocks affect scion nutrition, fruit quality, growth, yield and economical return. **Fruits**, v.55, n.4, p.231–239, 2000. Available from: <<https://journals.ashs.org/hortsci/view/journals/hortsci/35/3/article-p499C.xml>>. Accessed: Apr. 15, 2021. doi: 10.21273/HORTSCI.35.3.499C.