RODRIGO SILVA ALVES

OPTIMIZATION OF FOREST TREE BREEDING THROUGH MULTIPLE-TRAIT MIXED MODELS

Thesis presented to the Universidade Federal de Viçosa as part of the requirements of the Graduate Program in Genetics and Breeding to obtain the title of Doctor of Science.

Advisor: Marcos Deon Vilela de Resende

VIÇOSA - MINAS GERAIS
2020

Ficha catalográfica preparada pela Biblioteca Central da Universidade Federal de Viçosa - Câmpus Viçosa

T

A474o
2020

Alves, Rodrigo Silva, 1988-
Optimization of forest tree breeding through multiple-trait mixed models / Rodrigo Silva Alves. – Viçosa, MG, 2020.
49 f. : il. (algumas color.) ; 29 cm.

Texto em inglês.
Inclui anexos.
Orientador: Marcos Deon Vilela de Resende.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. BLUP. 2. Análise de variância. 3. Modelos multiníveis (Estatísticas). 4. Florestas - Melhoramento genético. I. Universidade Federal de Viçosa. Departamento de Biologia Geral. Programa de Pós-Graduação em Genética e Melhoramento. II. Título.

CDD 22. ed. 576.5

RODRIGO SILVA ALVES

OPTIMIZATION OF FOREST TREE BREEDING THROUGH MULTIPLE-TRAIT MIXED MODELS

Thesis presented to the Universidade Federal de Viçosa as part of the requirements of the Graduate Program in Genetics and Breeding to obtain the title of Doctor of Science.

APPROVED: February 28th, 2020.

Assent:

Rodrigo S. Alves

Rodrigo Silva Alves
Author

MResende

Marcos Deon Vilela de Resende
Advisor

To my parents, Bráz Tadeu Alves and Maria Lúcia Silva Alves, for their unconditional effort to get me here.

DEDICATION

ACKNOWLEDGMENTS

To God, for giving me life and providing everything I am and own.

To my parents, Bráz Tadeu Alves and Maria Lúcia Silva Alves, for their love, support and encouragement. To my brother, Luciano Silva Alves, for the companionship.

To Professor Marcos Deon Vilela de Resende, for his guidance, trust, friendship and the professionalism for which he should be looked up to.

To Professor Leonardo Lopes Bhering, for the partnership and trust.

To Researchers Bruno Galveas Laviola and Eduardo Pinheiro Henriques, for providing the experimental data.

To Professors Antônio Policarpo Souza Carneiro, Fabyano Fonseca e Silva, Luiz Antônio dos Santos Dias and Rogério Luiz da Silva, for their contribution to the improvement of this work.

To my friends at the Biometrics Laboratory, especially João Romero do Amaral Santos de Carvalho Rocha and Paulo Eduardo Teodoro.

To the Universidade Federal de Viçosa and the Graduate Program in Genetics and Breeding, for their excellence and for the opportunity given to me.

To the Coordination for the Improvement of Higher Education Personnel (CAPES), for the Doctoral Fellowship grant.

To everyone who directly or indirectly contributed to my professional qualification.

THANK YOU VERY MUCH!

The present work was financially supported by the Coordination for the Improvement of Higher Education Personnel (CAPES) - Finance Code 001.

RESUMO

ALVES, Rodrigo Silva, D.Sc., Universidade Federal de Viçosa, fevereiro de 2020. **Otimização do melhoramento florestal por meio de modelos mistos multicaracterísticos**. Orientador: Marcos Deon Vilela de Resende.

As técnicas atualmente usadas para avaliação genética envolvem a predição de valores genéticos e a estimação de componentes de variância simultaneamente. O método amplamente utilizado, em todo o mundo, para predição de valores genéticos é a melhor predição linear não viesada (BLUP). A predição usando o BLUP assume que os componentes de variância são conhecidos. No entanto, na prática, esses valores não são conhecidos e devem ser estimados com precisão para obter o BLUP empírico. Atualmente, o método padrão para estimação de componentes de variância é o da máxima verossimilhança restrita (REML). Portanto, a avaliação genética consiste em usar essas metodologias juntas (procedimento REML/BLUP). Porém, quando a seleção genética é baseada em várias características, que podem ser geneticamente correlacionadas devido a genes pleiotrópicos e/ou desequilíbrio da fase gamética, pode ocorrer viés de seleção se essas características forem analisadas individualmente. Assim, o presente trabalho teve como objetivo avaliar a aplicabilidade e eficiência do BLUP multicaracterístico no melhoramento florestal. Os modelos multicaracterísticos empregados conduziram a melhores resultados que os tradicionais modelos univariados e de repetibilidade. Portanto, o BLUP multicaracterístico pode ser utilizado com vantagem no melhoramento florestal.

Palavras-chave: BLUP. REML. Metodologia de modelos mistos.

ABSTRACT

ALVES, Rodrigo Silva, D.Sc., Universidade Federal de Viçosa, February, 2020. **Optimization of forest tree breeding through multiple-trait mixed models**. Advisor: Marcos Deon Vilela de Resende.

The current techniques used for genetic evaluation involve the simultaneous prediction of genetic values and estimation of variance components. Best linear unbiased prediction (BLUP) is a widely used method worldwide to predict genetic values. Prediction using BLUP assumes that the variance components are known. However, these values are not known in practice and should thus be accurately estimated to obtain the empirical BLUP. At present, the standard method to estimate variance components is restricted maximum likelihood (REML). Therefore, genetic evaluation consists of using these methodologies together (REML/BLUP procedure). Nonetheless, when genetic selection is based on several traits, which may be genetically correlated due to pleiotropic genes and/or gametic phase imbalance, selection bias may occur if these traits are analyzed individually. Thus, the present work aimed to evaluate the applicability and efficiency of multiple-trait BLUP in forest-tree breeding. The multiple-trait models employed led to better results than the traditional univariate and repeatability models. On this basis, the multiple-trait BLUP can be advantageously used in forest breeding.

Keywords: BLUP. REML. Mixed model methodology.

CONTENTS

GENERAL INTRODUCTION

In stricter terms, genetic selection is a purely statistical problem, since, in practice, a fraction of individuals is selected according to their genetic values, which follow a probability distribution (RESENDE, 2007a). A statistical model refers to a mathematical expression that describes the behavior of a random variable as a function of other classificatory or explanatory variables. Linear models are those linear in their parameters, that is, the parameters that integrate the models are simple coefficients of the classificatory or explanatory variables (RESENDE *et al.*, 2014). Linear models are described in detail by Henderson (1984) and Searle (1987).

Henderson (1949) developed the best linear unbiased prediction method (BLUP), through which fixed effects are estimated and random effects are predicted simultaneously. The properties of BLUP are similar to those of a selection index and the method is reduced to selection indices when no adjustments for environmental factors are necessary (MRODE, 2014). The properties of BLUP are more or less incorporated in its name (RESENDE *et al.*, 2014):

Best: minimizes prediction-error variance (PEV), i.e., maximizes precision;

Linear: is a linear function of observations;

Unbiased: is not biased, which is a property that, along with the minimization of PEV, maximizes accuracy in the class of unbiased predictors; and

Prediction: predicts a random variable.

The BLUP method was initially applied in the animal breeding. Its use was maximized with the increasing computational efficiency and it evolved from simple models such as the parent model, in its early years, to more complex models such as the individual, multiple-trait and random regression models (MRODE, 2014). In the 1990s, a few specialized software programs were developed in which the BLUP procedure was implemented, with Selegen REML/BLUP (RESENDE, 2007b; RESENDE, 2016) and ASReml (GILMOUR *et al.*, 2015a; GILMOUR *et al.*, 2015b) standing out for plant breeding.

Genetic selection is usually based on a combination of several traits of economic importance that may or may not be genetically correlated (RESENDE, 2015). Such traits can be combined into an index in which individuals are classified. Multiple-trait analysis is the ideal procedure to evaluate individuals in these situations, as it explains the interrelationships between the traits (MRODE, 2014). In this respect, the multiple-trait BLUP, developed by Henderson and Quaas (1976), allows the evaluation of individuals for two or more traits simultaneously, using genetic and residual correlations.

One of the main advantages of multiple-trait BLUP is that it increases selection accuracy. The gain in precision depends on the absolute difference between the genetic and residual correlations between the traits. The greater the differences in these correlations, the greater the increase in accuracy (SCHAEFFER, 1984; THOMPSON, MEYER, 1986). When, for example, heritability and genetic and residual correlations for two traits are the same, the multiple-trait predictions are equivalent to those of single-trait analyses. Moreover, traits of lower heritability benefit more when analyzed with traits of higher heritability in multiple-trait analysis. Another advantage is an additional increase in accuracy resulting from the better connections between the data due to the residual covariance between traits (THOMPSON, MEYER, 1986).

In some cases, only one trait is used in early selection. For instance, only plants resistant to a given pathogen are selected and subsequently evaluated for the yield trait. In this case, a single-trait analysis of yield can be biased, as it does not include information on the resistance trait on which early selection was based. This is often called *selection bias*. However, a multiple-trait analysis considering the resistance and yield traits can eliminate this bias. Thus, multiple-trait BLUP prevents the occurrence of selection bias.

One of the disadvantages of multiple-trait analysis is the high computational cost. The cost of multiple analysis of *n* traits is much higher than the cost of *n* single-trait analyses. Another disadvantage of the former approach is that it requires reliable estimates of genetic and phenotypic correlations between traits, which may not be readily available (MRODE, 2014).

The prediction of genetic values depends, essentially, on reliable estimates of variance components, which are assumed to be known in the BLUP procedure. However, in practice, they must be estimated with the greatest possible precision and accuracy (RESENDE, 2002).

The restricted maximum likelihood method (REML), developed by Patterson and Thompson (1971), is the ideal procedure for estimating variance components with balanced or unbalanced data. The REML method requires an iterative solution due to the non-linearity of the equations, which makes it difficult to derive explicit estimators. Thus, the variance components iterate in the mixed-model BLUP equations until convergence (RESENDE, 2002). Most often, variance components are estimated and genetic values are predicted simultaneously. In this case, the BLUP procedure is called *empirical BLUP*.

Under certain conditions, maximum likelihood methods can take the effects of selection into account. All the information that contributed to the selection must be included in the analysis (as well as the relationship between individuals), except if this information is not correlated with the trait under analysis (RESENDE, 2002). Even if these conditions are only partially met, these methods provide less biased estimates than those obtained in methods based on analysis of variance (MEYER, 1989). In general, theoretical arguments and evidence indicate that specific inferences made from likelihood functions are not affected by some forms of selection (GIANOLA *et al*., 1989). This property rendered REML the standard method for estimating variance components in breeding programs.

The REML method eliminates bias due to changes in allele frequencies through selection, using the full kinship matrix (A, G or H) (RESENDE, 2015). This makes it is possible to obtain variance components for an unselected base population and predict genetic values of individuals of any generation with precision. The use of the full kinship matrix takes into account changes in genetic variance due to inbreeding and linkage disequilibrium, which stem from selection (KENNEDY, SORENSEN, 1988), in addition to considering the genetic trend or genetic gain achieved (RESENDE, 2002).

The REML method requires normality so that the estimators have desirable properties. However, such estimators may be robust to deviations from normality, generating reasonable estimates even when the form of distribution is not specified (HARVILLE, 1977; RESENDE, 2002).

For balanced data, the REML estimator (neglecting the normality assumption and the restriction of REML estimates to the parametric space) is identical to the least squares estimator (analysis of variance), which presents the desirable BUE (best unbiased estimator) and BQUE (best quadratic unbiased estimator) properties under normality (SEARLE *et al*., 1992; RESENDE, 2002).

The REML/BLUP method is a flexible and powerful tool to estimate variance components and fixed effects and predict random effects such as genetic values and statistical analysis in general (RESENDE, 2002). The main practical advantages of the REML/BLUP procedure (PIRES *et al*., 2011; RESENDE, 2016) are that it allows a comparison of individuals or varieties across time (generations, years) and space (locations, blocks); simultaneous correction for environmental effects, estimation of variance components and prediction of genetic values; handling complex data structures (repeated measures, different years, locations and designs); and can be applied to unbalanced data and non-orthogonal designs.

Therefore, the present study proposes to approach modeling, inference, estimation of variance components and prediction of genetic values associated with continuous variables using the multiple-trait REML/BLUP method.

REFERENCES

GIANOLA, D.; FERNANDO, R. L.; IM, S.; FOULLEY, J. L. Likelihood estimation of quantitative genetic parameters when selection occurs: models and problems. **Genome**, v. 31, p. 768-777, 1989.

GILMOUR, A. R.; GOGEL, B. J.; CULLIS, B. R.; WELHAM, S. J.; THOMPSON, R. **ASReml User Guide Release 4.1 Functional Specification**. Hemel Hempstead: VSN International, 2015a. 364p.

GILMOUR, A. R.; GOGEL, B. J.; CULLIS, B. R.; WELHAM, S. J.; THOMPSON, R. **ASReml User Guide Release 4.1 Structural Specification**. Hemel Hempstead: VSN International, 2015b. 375p.

HARVILLE, D. A. Maximum likelihood approaches to variance component estimation and to related problems. **Journal of the American Statistical Association**, v. 72, p. 320-328, 1977.

HENDERSON, C. R. **Applications of linear models in animal breeding**. Guelph: University of Guelph, 1984. 462p.

HENDERSON, C. R. Estimation of changes in herd environment. **Journal of dairy science**, v. 32 p. 709, 1949.

HENDERSON, C. R.; QUAAS, R. L. Multiple trait evaluation using relatives records. **Journal of Animal Science**, v. 43 n. 6, p. 1188-1197, 1976.

KENNEDY, B. W.; SORENSEN, D. A. Properties of mixed model methods for prediction of genetic merit under different genetic models in selected and unselected populations. In: WEIR, B.; GOODMAN, M. M.; NAMKOONG, G. (Ed.). **Second international Conference on Quantitative Genetics**. Raleigh: North Carolina State University, 1988. p. 91-103.

MEYER, K. Estimation of genetic parameters. In: HILL, W. G.; MACKAY, T. F. C. **Evolution and animal breeding**. Wallingford: CAB International, 1989. p. 161-167.

MRODE, R. A. **Linear models for the prediction of animal breeding values**. Wallingford: CAB International, 2014. 360p.

PATTERSON, H. D.; THOMPSON, R. Recovery of inter-block information when block sizes are unequal. **Biometrika**, v. 58, n. 3, p. 545-554, 1971.

PIRES, I. E.; RESENDE, M. D. V.; SILVA, R. L.; RESENDE JR., M. F. R. **Genética Florestal**. Viçosa: Arka, 2011. 318p.

RESENDE, M. D. V. **Genética Biométrica e Estatística no Melhoramento de Plantas Perenes**. Brasília: Embrapa Informação Tecnológica, 2002. 975p.

RESENDE, M. D. V. **Genética Quantitativa e de Populações**. Visconde do Rio Branco: Suprema, 2015. 422p.

RESENDE, M. D. V. **SELEGEN-REML/BLUP: Sistema Estatístico e Seleção Genética Computadorizada via Modelos Lineares Mistos**. Colombo: Embrapa Florestas, 2007b. 360p.

RESENDE, M. D. V. Software Selegen-REML/BLUP: a useful tool for plant breeding. **Crop Breeding and Applied Biotechnology**, v 16, n. 4, p. 330-339.

RESENDE, M. D. V.; SILVA, F. F.; AZEVEDO, C. F. **Estatística Matemática, Biométrica e Computacional: Modelos Mistos, Multivariados, Categóricos e Generalizados (REML/BLUP), Inferência Bayesiana, Regressão Aleatória, Seleção Genômica, QTL-GWAS, Estatística Espacial e Temporal, Competição, Sobrevivência**. Visconde do Rio Branco: Suprema, 2014. 881p.

RESENDE, M. D. V. **Matemática e Estatística na Análise de Experimentos e no Melhoramento Genético**. Colombo: Embrapa Florestas, 2007a. 561p.

SCHAEFFER, L. R. Sire and cow evaluation under multiple trait models. **Journal of Dairy Science**, v. 67, p. 1567-1580, 1984.

SEARLE, S. R. **Linear models for unbalanced data**. New York: Wiley-Blackwell, 1987. 560p.

SEARLE, S. R.; CASELLA, G.; Mc CULLOCH, C. E. **Variance components**. New York: John Wiley & Sons, 1992. 536p.

THOMPSON, R.; MEYER, K. A review of theoretical aspects in the estimation of breeding values for multi-trait selection. **Livestock Production Science**, v. 15, p. 299-313, 1986.

CHAPTER 1

MULTIPLE-TRAIT BLUP IN THE GENETIC SELECTION OF *EUCALYPTUS* FAMILIES

VIÇOSA - MINAS GERAIS

2020

ABSTRACT

Usually, genetic selection is carried out based on several traits, which can be genetically correlated. In this case, selection bias may occur if these traits are analyzed individually. Thus, the present work aimed to evaluate the applicability and efficiency of multiple-trait best linear unbiased prediction (BLUP) in the genetic selection of *Eucalyptus* families. The data used in this work refer to the evaluation of a partial diallel of *Eucalyptus* spp. in relation to height, diameter at breast height (DBH), and volume. Variance components and genetic and non-genetic parameters were estimated via residual maximum likelihood (REML). Multiple-trait BLUP led to estimates of mean additive genetic variance higher than the estimates obtained via single-trait BLUP and, consequently, led to higher estimates of narrow-sense individual interpopulational heritabilities; and mean accuracies. Partial genetic correlations obtained via multiple-trait BLUP allowed a real understanding of the association between traits, differently from those obtained via single-trait BLUP. Multiple-trait BLUP led to higher gains predicted with the selection for height, DBH, and volume; and can be efficiently applied in the genetic selection of *Eucalyptus* families.

Key words: *Eucalyptus* breeding, mixed model methodology, diallel, genetic correlation

1 INTRODUCTION

The genus *Eucalyptus* is one of the main forest species cultivated in tropical and subtropical regions. Its cultivation is mainly intended to produce pulp, bio-oil, firewood, and charcoal. The international pulp trade and the intense search for alternative energy sources have increasingly motivated the establishment of *Eucalyptus* plantations in several countries worldwide (Fonseca et al. 2010). In this sense, *Eucalyptus* breeding programs have sought to identify more efficient selection methods to increase yield and quality of traits of industrial interest.

The current techniques used for genetic evaluation involve variance components estimation and genetic values prediction simultaneously (Resende 2016). The standard method for estimating variance components is the residual maximum likelihood (REML), developed by Patterson and Thompson (1971). The traditional method used to predict genetic values is the best linear unbiased prediction (BLUP), formally presented by Henderson (1975). Therefore, the optimum procedure for genetic evaluation is the REML/BLUP, also defined as mixed model methodology.

The main advantages of using mixed model methodology are the fact that they allow incorporating kinship information; comparing individuals or varieties over time and space; correcting environmental effects, variance components estimate, and genetic values prediction simultaneously; and dealing with complex data structures. Moreover, the mixed model methodology may be applied to unbalanced data and not orthogonal designs (Mrode 2014; Resende 2016).

The single-trait BLUP has been commonly used in plant breeding (Resende 2016), providing good results for several crops, including *Eucalyptus* (Alves et al. 2018; Corrêa et al. 2017; Costa et al. 2016; Resende et al. 2016; Zhu et al. 2017). However, when genetic selection is based on several traits, which may be genetically correlated due to pleiotropic genes and/or gametic phase imbalance (Falconer and Mackay 1996), selection bias may occur if these traits are analyzed individually (Pollak et al. 1984).

In this sense, to eliminate the selection bias and increase predictive accuracy, Henderson and Quaas (1976) proposed the multiple-trait BLUP. This method tends to be more efficient since it considers a greater amount of data and uses the genetic and residual correlations between traits. The multiple-trait BLUP is commonly used in animal breeding (Mrode 2014) and has been successfully applied in plant breeding (Bauer and Léon 2008; Costa et al. 2002; Imai et al. 2016; Piepho et al. 2008; Viana et al. 2010). Therefore, this work aimed to evaluate the applicability and efficiency of multiple-trait BLUP in the genetic selection of *Eucalyptus* families.

2 MATERIALS AND METHODS

2.1 EXPERIMENTAL DATA

The data used in this work refer to the evaluation of a partial diallel of *Eucalyptus* spp., whose female parent population was composed of 33 trees selected based on a progeny test of *E. urophylla* and male parent population was composed of 64 trees selected in clonal banks, clonal tests, and seed orchards, based on their volumetric yields and species variability and hybrids (Table S1).

The experiment was implemented in November 2003, in a randomized block design, with eight replications and six plants per plot, arranged in rows, spaced 3 m between rows and 2 m between plants. It was settled in the experimental field of Aperam BioEnergia Company, Itamarandiba, Minas Gerais, Brazil (17°44'45" S and 42°45'11" W; 1000 m asl). The experiment consisted of the performance evaluation of 11458 7-year-old individuals from 286 full-sib families, regarding height (m), diameter at breast height (DBH) (cm), and volume (m^3). Volume was obtained based on the following expression:

$$Volume = (\pi DBH^2/40000) * height * \bar{F},$$

where: $\bar{F}$ is the average form factor obtained for the population ($\bar{F} = 0.5$).

2.2 STATISTICAL ANALYSES

The mixed model methodology was applied to estimate the variance components and predict the genetic values via single- and multiple-trait BLUP, according to Gilmour et al. (2015). The statistical model associated with the evaluation of full-sib families obtained under diallel crosses was determined by the following equation:

$$y = Xr + Ts + Wf + Zm + Td + Qp + e,$$

where: *y* is the vector of phenotypes, *r* is the vector of fixed effects of replication added to the general mean, *s* is the vector of fixed effects of species, *f* is the vector of the random effects of female parents, *m* is the vector of the random effects of male parents, *d* is the vector of the random effects of families (specific combining ability or dominance), *p* is the vector of random effects of plot, and *e* is the vector of random residues. Uppercase letters represent the incidence matrices for these effects.

In the single-trait BLUP: $f \sim N(0, \sigma^2_{fem})$, $m \sim N(0, \sigma^2_{mal})$, $d \sim N(0, \sigma^2_{sca})$, $p \sim N(0, \sigma^2_{plot})$, and $e \sim N(0, \sigma^2_e)$; where: σ^2_{fem} is the genetic variance between female parents, σ^2_{mal} is the genetic variance between male parents, σ^2_{sca} is the variance of specific combining ability between two parents, σ^2_{plot} is the variance between plots, and σ^2_e is the residual variance.

In the multiple-trait BLUP: $f \sim N(0, \sum_{fem} \otimes I)$, $m \sim N(0, \sum_{mal} \otimes I)$, $d \sim N(0, \sum_{sca} \otimes I)$, $p \sim N(0, \sum_{plot} \otimes I)$, and $e \sim N(0, \sum_e \otimes I)$; where: $\sum_{fem}$ is the genetic covariance matrix between female parents, $\sum_{mal}$ is the genetic covariance matrix between male parents, $\sum_{sca}$ is the covariance matrix of specific combining ability between two parents, $\sum_{plot}$ is the covariance matrix between plots, $\sum_e$ is the residual covariance matrix, I is an identity matrix with appropriate order to the respective random effect, and $\otimes$ denotes the Kronecker product. In addition, $\sum_{fem}$, $\sum_{mal}$, $\sum_{sca}$, $\sum_{plot}$, and $\sum_e$ are unstructured covariance structures.

Mean additive genetic variance (σ_a^2), phenotypic variance (σ_{phen}^2), narrow-sense individual heritability in the female parents population ($h_{a\,fem}^2$), narrow-sense individual heritability in the male parents population ($h_{a\,mal}^2$), coefficient of determination of the effects of the specific combining ability (c_{sca}^2), individual heritability of interpopulational dominance effects (h_{dom}^2), narrow-sense individual interpopulational heritability (h_a^2), broad-sense individual interpopulational heritability (h_g^2), and coefficient of determination of plot effects (c_{plot}^2) were obtained, respectively, by the following expressions:

$$\sigma_a^2 = 2\sigma_{fem}^2 + 2\sigma_{mal}^2,$$

$$\sigma_{phen}^2 = \sigma_{fem}^2 + \sigma_{mal}^2 + \sigma_{sca}^2 + \sigma_{plot}^2 + \sigma_e^2,$$

$$h_{a\,fem}^2 = 4\sigma_{fem}^2/\sigma_{phen}^2,$$

$$h_{a\,mal}^2 = 4\sigma_{mal}^2/\sigma_{phen}^2,$$

$$c_{sca}^2 = \sigma_{sca}^2/\sigma_{phen}^2,$$

$$h_{dom}^2 = 4\sigma_{sca}^2/\sigma_{phen}^2,$$

$$h_a^2 = \sigma_a^2/\sigma_{phen}^2,$$

$$h_g^2 = (\sigma_a^2 + 4\sigma_{sca}^2)/\sigma_{phen}^2, \text{ and}$$

$$c_{plot}^2 = \sigma_{plot}^2/\sigma_{phen}^2.$$

The standard errors of the estimates of variance components and of the genetic and non-genetic parameters were obtained through mixed model output and post-processing "pin" function of the ASReml software.

Mean accuracy of female parents ($r_{\hat{a}a\,fem}$), of male parents ($r_{\hat{a}a\,mal}$), and of families ($r_{\hat{a}a\,sca}$), were obtained by the following expression (Resende et al. 2014):

$$r_{\hat{a}a\,x} = \left[\sum_{i=1}^{n}\sqrt{1-(PEV/\sigma_x^2)}\right]/n,$$

where: *x* refers to female parents (*fem*) or male parents (*mal*) or families (*sca*), *n* is the number of female parents or male parents or families; and *PEV* is the variance of the prediction error, associated to *x*, extracted from the diagonal of generalized inverse of the coefficient matrix of the mixed model equations.

To obtain the genetic covariances via single-trait BLUP, the analyses of the sum of the phenotypic values - in pairs - of the traits were performed. Thus, the genetic covariance between traits i and j ($\sigma_{g_{i,j}}$) were obtained based on the following expression:

$$\sigma_{g_{i,j}} = (\sigma^2_{g_{i+j}} - \sigma^2_{g_i} - \sigma^2_{g_j})/2$$

where: $\sigma^2_{g_{i+j}}$, $\sigma^2_{g_i}$ and $\sigma^2_{g_j}$ are the genetic variances of traits *i plus j*, *i*, and *j*, respectively. For the multiple-trait BLUP, the genetic covariances were obtained directly by mixed model output.

The simple genetic correlation (Pearson correlation) between traits i and j ($r_{i,j}$), via single- and multiple-trait BLUP, were obtained based on the following expression:

$$r_{i,j} = \sigma_{g_{(i,j)}} \Big/ \sqrt{\sigma^2_{g_i} \sigma^2_{g_j}}.$$

Was verified if the matrices of simple genetic correlation are positive semidefinite. In negative case, the transformation proposed by Schaeffer (2016) was used to make it positive semidefinite. The partial genetic correlation between traits i and j eliminating statistically the influence of trait k ($r_{ij.k}$), via single- and multiple-trait BLUP, were obtained through the inverse of simple genetic correlation matrix, based on the following expression:

$$r_{ij.k} = -a_{ij}/\sqrt{a_{ii}a_{jj}},$$

where: a_{ij}, a_{ii}, and a_{jj} are, respectively, the elements of order *ij*, *ii*, and *jj* of the inverse of the simple genetic correlation matrix.

The rank correlations (Spearman correlation) (ρ) of the predicted genetic values for height, DBH, and volume by both methods, for families, were given by:

$$\rho = 1 - [6\textstyle\sum_{i=1}^{n} d_i^2/(n^3 - n)],$$

where: d_i is the difference between the two ranks of each family, and n is the number of families.

The agreement between selected families, via single- and multiple-trait BLUP, was calculated using the Cohen's kappa coefficient (*K*) (Cohen 1960) given by:

$$K = [(A - C)/(D - C)]100,$$

where: A is the number of coincident families by the two methods, D is the number of families selected, and C is the number of families coincident due to chance ($C = bD$, where: b is the selection intensity).

Selection gains (*SG*), for each trait, was predicted, considering selection intensities equal to 10%, 15%, and 20%, based on the following expression:

$$SG = (\sum_{i=1}^{n} GV_i)/n,$$

where: GV_i is the predicted genetic value of family *i*, and *n* is the number of families selected.

Statistical analyses were performed using the ASReml 4.1 (Gilmour et al. 2015) and R (R Development Core Team 2018) software.

3 RESULTS

3.1 VARIANCE COMPONENTS

The standard errors of the estimates of the variance components revealed that all random effects of the statistical model are statistically significant (i.e., they differ from zero, with a confidence level of 66%) (Table 1). As expected, given that the female parent population was composed by a single species and the male parent population was composed by many species, the estimates of genetic variance between male parents were higher than the estimates of genetic variance between female parents and, consequently, contributed with the greatest fraction of the mean additive genetic variance, regardless of the evaluated trait and method used (Table 1).

Table 1. Estimates of variance components and their respective standard error, between parenthesis, for the traits: height, diameter at breast height (DBH), and volume, evaluated in 286 full-sib families of *Eucalyptus*, via single- and multiple-trait BLUP.

Component	Single-trait BLUP			Multiple-trait BLUP		
	Height	DBH	Volume	Height	DBH	Volume
σ^2_{fem}	0.3231 (0.1299)	0.2138 (0.0844)	0.0003 (0.0001)	0.3280 (0.1302)	0.2266 (0.0872)	0.0004 (0.0001)
σ^2_{mal}	0.6200 (0.2496)	0.4177 (0.1650)	0.0010 (0.0003)	0.6266 (0.2408)	0.4694 (0.1666)	0.0010 (0.0003)
σ^2_{a}	1.8863 (0.5632)	1.2631 (0.3712)	0.0026 (0.0007)	1.9093 (0.5466)	1.3919 (0.3760)	0.0026 (0.0007)
σ^2_{sca}	0.6227 (0.1347)	0.3812 (0.0805)	0.0005 (0.0001)	0.6160 (0.1327)	0.3675 (0.0778)	0.0005 (0.0001)
σ^2_{plot}	0.9563 (0.1796)	0.1092 (0.1030)	0.0003 (0.0002)	0.9533 (0.1793)	0.1103 (0.1030)	0.0003 (0.0002)
σ^2_{e}	21.7836 (0.3207)	14.5880 (0.2144)	0.0208 (0.0003)	21.7912 (0.3207)	14.5887 (0.2144)	0.0208 (0.0003)
σ^2_{phen}	24.3058 (0.3306)	15.7099 (0.2070)	0.0229 (0.0004)	24.3151 (0.4046)	15.7624 (0.2680)	0.0229 (0.0004)

σ^2_{fem}: genetic variance between female parents, σ^2_{mal}: genetic variance between male parents, σ^2_{a}: mean additive genetic variance, σ^2_{sca}: variance of the specific combining ability between two parents, σ^2_{plot}: variance between plots, σ^2_{e}: residual variance, σ^2_{phen}: phenotypic variance.

Multiple-trait BLUP led to estimates of mean additive genetic variance higher than the estimates obtained via single-trait BLUP (Table 1) and, consequently, led to higher estimates of narrow-sense individual interpopulational heritabilities; and mean accuracy of female parents, male parents, and families (Table 2). The estimates of mean additive genetic variance presented increments of 1.22%, 10.20%, and 1.54%; the estimates of narrow-sense individual interpopulational heritability presented increments of 1.16%, 9.83%, and 1.41%; and the estimates of mean accuracy of families presented increments of 3.65%, 1.05%, and 2.08%, for height, DBH, and volume, respectively.

Table 2. Estimates of genetic and non-genetic parameters and their respective standard error, between parenthesis, for the traits: height, diameter at breast height (DBH), and volume, evaluated in 286 full-sib families of *Eucalyptus*, via single- and multiple-trait BLUP.

Parameter	Single-trait BLUP			Multiple-trait BLUP		
	Height	DBH	Volume	Height	DBH	Volume
$h^2_{a\,fem}$	0.0532	0.0544	0.0592	0.0540	0.0575	0.0615
	(0.0211)	(0.0212)	(0.0225)	(0.0212)	(0.0219)	(0.0231)
$h^2_{a\,mal}$	0.1020	0.1064	0.1669	0.1031	0.1191	0.1679
	(0.0402)	(0.0411)	(0.0543)	(0.0387)	(0.0412)	(0.0528)
c^2_{sca}	0.0256	0.0243	0.0237	0.0253	0.0233	0.0236
	(0.0055)	(0.0051)	(0.0051)	(0.0054)	(0.0049)	(0.0050)
h^2_{dom}	0.1025	0.0971	0.0948	0.1013	0.0932	0.0943
	(0.0221)	(0.0205)	(0.0204)	(0.0217)	(0.0197)	(0.0201)
h^2_a	0.0776	0.0804	0.1131	0.0785	0.0883	0.1147
	(0.0224)	(0.0228)	(0.0290)	(0.0217)	(0.0229)	(0.0283)
h^2_g	0.1801	0.1775	0.2078	0.1799	0.1816	0.2089
	(0.0262)	(0.0252)	(0.0298)	(0.0261)	(0.0258)	(0.0297)
c^2_{plot}	0.0393	0.0069	0.0128	0.0392	0.0070	0.0128
	(0.0073)	(0.0066)	(0.0066)	(0.0073)	(0.0065)	(0.0066)
$r_{\hat{a}a\,fem}$	0.7313	0.7471	0.7581	0.7386	0.7636	0.7671
$r_{\hat{a}a\,mal}$	0.5986	0.6172	0.6669	0.6482	0.6879	0.6922
$r_{\hat{a}a\,sca}$	0.6088	0.6219	0.6049	0.6310	0.6284	0.6175

$h^2_{a\,fem}$: narrow-sense individual heritability in the female parents population, $h^2_{a\,mal}$: narrow-sense individual heritability in the male parents population, c^2_{sca}: coefficient of determination of the effects of the specific combining ability, h^2_{dom}: individual heritability of interpopulational dominance effects, h^2_a: narrow-sense individual interpopulational heritability, h^2_g: broad-sense individual interpopulational heritability, c^2_{plot}: coefficient of determination of plot effects, $r_{\hat{a}a\,fem}$: mean accuracy of female parents, $r_{\hat{a}a\,mal}$: mean accuracy of male parents, and $r_{\hat{a}a\,sca}$: mean accuracy of families.

Conversely, for all traits, estimates of the variance of the specific combining ability between two parents (Table 1) and individual heritability of the interpopulational dominance effects (Table 2) obtained via single-trait BLUP were slightly higher.

3.2 GENETIC INTERRELATION OF TRAITS

It was verified that the simple genetic correlation matrix obtained via single-trait BLUP for male parents isn't a positive semidefinite matrix (negative determinant). In addition, the simple genetic correlation between DBH and volume, exceeded the parametric space (Table 3). This bias did not occur when the simple genetic correlation between this pair of traits was estimated via multiple-trait BLUP.

Table 3. Simple genetic correlations between the traits: height, diameter at breast height (DBH), and volume, evaluated in 286 full-sib families of *Eucalyptus* via single-trait BLUP (above the diagonal) and multiple-trait BLUP (below the diagonal).

Trait	Height	DBH	Volume
Height	-	0.86 [a] | 0.88 [b] | 0.98 [c]	0.85 [a] | 0.82 [b] | 0.95 [c]
DBH	0.86 [a] | 0.89 [b] | 0.97 [c]	-	0.98 [a] | 1.02 [b] | 0.98 [c]
Volume	0.85 [a] | 0.87 [b] | 0.93 [c]	0.99 [a] | 1.00 [b] | 0.98 [c]	-

[a] female parent, [b] male parent, and [c] families.

For female parents, signal inversion of the partial genetic correlations occurred, between volume and DBH, when obtained via single- and multiple-trait BLUP (Table 4). For male parents and families, no signal inversion of the partial genetic correlations obtained via single- and multiple-trait BLUP was observed; however, differences in magnitudes were detected. For the male parents, the magnitudes of partial genetic correlations between all the traits reduced when obtained via multiple-trait BLUP. Conversely, for the families, the magnitudes of partial genetic correlations increased when obtained via multiple-trait BLUP, regardless of the traits.

Table 4. Partial genetic correlations between the traits: height, diameter at breast height (DBH), and volume, evaluated in 286 full-sib families of *Eucalyptus* via single-trait BLUP (above the diagonal) and multiple-trait BLUP (below the diagonal).

Trait	Height	DBH	Volume
Height	-	0.27 [a] | 0.99 [b] | 0.79 [c]	0.03 [a] | -0.99 [b] | -0.17 [c]
DBH	0.30 [a] | 0.60 [b] | 0.85 [c]	-	0.94 [a] | 1.00 [b] | 0.72 [c]
Volume	-0.03 [a] | -0.52 [b] | -0.54 [c]	0.95 [a] | 0.99 [b] | 0.89 [c]	-

[a] female parent, [b] male parent, and [c] families.

3.3 FAMILIES SELECTION

The single- and multiple-trait BLUP were highly correlated in terms of ranking of families ($\rho = 0.96$, $\rho = 0.94$, and $\rho = 0.98$ for height, DBH, and volume, respectively). In general, both methods presented a substantial agreement between the families selected (Table 5). Besides that, it was observed that the agreement between selected families increased with the increase in the selection intensity for height and DBH (Table 5).

Table 5. Selection intensity, selected quantity, and Cohen's kappa coefficient between selected families via single- and multiple-trait-BLUP for the traits: height, diameter at breast height (DBH), and volume evaluated in 286 full-sib families of *Eucalyptus*.

Selection intensity (%)	Selected quantity	Cohen's kappa coefficient (%)		
		Height	DBH	Volume
10	29	73.18	61.69	73.18
15	43	75.38	72.64	89.06
20	57	80.26	73.68	84.65

For all traits, multiple-trait BLUP led to higher gains predicted with the selection than single-trait BLUP (Table 6). Height, DBH and volume presented, respectively, the following increases when obtained via multiple-trait BLUP: 5.73%, 8.18%, and 3.35% at a selection intensity of 10%; 5.10%, 8.40%, and 1.38% at a selection intensity of 15%; and 3.27%, 7.34%, and 2.07% at a selection intensity of 20%. These results are consistent with the narrow-sense individual interpopulational heritability and mean accuracy of families (Table 2).

Table 6. Selection gains predicted in the unit of the trait and in percentage, between parenthesis, according to the selection intensity via single- and multiple-trait BLUP for the traits: height, diameter at breast height (DBH), and volume evaluated in 286 full-sib families of *Eucalyptus*.

Selection intensity	Single-trait BLUP			Multiple-trait BLUP		
(%)	Height	DBH	Volume	Height	DBH	Volume
10	0.7518 (3.19)	0.5816 (3.99)	0.0239 (10.68)	0.7949 (3.38)	0.6292 (4.32)	0.0247 (11.04)
15	0.6819 (2.90)	0.5179 (3.55)	0.0217 (9.68)	0.7167 (3.04)	0.5614 (3.85)	0.0220 (9.84)
20	0.6302 (2.68)	0.4757 (3.26)	0.0193 (8.62)	0.6508 (2.76)	0.5106 (3.50)	0.0197 (8.81)

4 DISCUSSION

4.1 VARIANCE COMPONENTS

The REML is a Fisherian method of modeling and inference but can be derived under the Bayesian approach, a fact that characterizes its generality as an optimal procedure. This method requires an iterative solution due to the nonlinearity of the equations, which makes it difficult to derive explicit estimators. Most of the time, the estimation of the variance components and the prediction of the genetic values are performed simultaneously by the REML/BLUP procedure (Resende et al. 2014). In this case, the BLUP procedure is denominated as empirical BLUP (Harville and Carriquiry 1992). Moreover, when BLUP is obtained for two or more traits simultaneously, it is denominated as multiple-trait BLUP, enabling more accurate estimates of the genetic and residual covariances between evaluated traits (Viana et al. 2010).

In this work, a diallel model (equivalent to the individual tree model) was used and its structure was preserved by using a reduced diallel individual model (similar to the reduced animal model). The greatest implementation difficulty relative to the individual tree model is the of solving the mixed model equations. Reduced diallel individual models capture the additive and dominance effects without an explicit kinship matrix (Resende et al. 2014). Therefore, diallel models are simpler alternative to the individual tree model.

In the identification and selection of superior genotypes, additive genetic variance is characterized as one of the major components to quantify the breeding potential of the study population, besides assisting in the selection strategy to be used to increase the frequency of favorable alleles (Kerr 1998; Resende 2015). In this work, the highest efficiency of the multiple-trait BLUP can be verified in the quantities estimated for mean additive genetic variance which reflected in higher values of narrow-sense individual interpopulational heritability and accuracies.

Despite the small differences, a close relationship was observed between the estimates for variance components and genetic and non-genetic parameters via single- and multiple-trait BLUP. Based on Viana et al. (2010), these results can be explained by the similarity in the value of the deviations between the estimation of genetic and residual variance components obtained via the single- and multiple-trait BLUP.

4.2 GENETIC INTERRELATION OF TRAITS

Falconer and Mackay (1996) assigned two main causes to genetic correlations: pleiotropy and genetic linkage. The latter is the most appropriate cause to explain the data due to the quantitative nature of the evaluated traits. Piepho et al. (2008) stated that for highly correlated traits, the multiple-trait BLUP is the most advantageous method. In the case where traits are analyzed individually (single-trait BLUP) selection bias may occur (Pollak et al. 1984) by leading to correlation values that extrapolate the parametric space (Viana et al. 2010). In this work, the simple genetic correlation between DBH and volume, obtained via single-trait BLUP, for male parents extrapolated the parametric space (Table 3). Besides that, the simple genetic correlation matrix obtained via single-trait BLUP for male parents isn't a positive semidefinite matrix and, therefore, required repair as proposed by Schaeffer (2016) to solve the problem.

Silveira et al. (2015) stated that the direct interpretation of simple genetic correlation between traits might result in mistakes in the selection strategies since it does not represent a cause and effect measurement. In this context, the partial genetic correlations have a notable use in plant breeding because it is a cause and effect measurement that gives the breeders the real relationship between pairs of traits (excluding the effect of other traits).

Inversion of signal was observed in the partial genetic correlations for female parents for the pair of traits: height and volume (Table 4). It is worth mentioning that the female parents were previously selected to form the diallel. However, the partial genetic correlation estimates obtained via multiple-trait BLUP are free of bias arising from previous selection (Henderson and Quaas 1976). In this way, the partial genetic correlation estimates obtained via multiple-trait BLUP allowed a real understanding of the association between height, DBH, and volume and therefore can be lead to the development of efficient selection strategies applied to *Eucalyptus* breeding.

4.3 FAMILIES SELECTION

In recent years, breeders have used the multiple-trait BLUP to predict genetic values in plant breeding. Persson and Andersson (2004) found that the prediction of genetic values via multiple-trait BLUP resulted in a lower mean bias and higher genetic correlation when compared with the single-trait BLUP in *Pinus sylvestris* L.. Costa et al. (2002) estimated genetic correlations in *Hevea brasiliensis* L. and revealed that multiple-trait BLUP is more accurate than single-trait BLUP. Imai et al. (2016) applied multiple-trait BLUP in citrus and demonstrated the usefulness of the method in predicting genetic values when the information of kinship among genotypes is known and in the presence of unbalanced data. Moreover, these authors recommended the multiple-trait BLUP as a tool to select genotypes.

The multiple-trait BLUP has also shown advantages over single-trait BLUP in different population structures. Bauer and Léon (2008) reported high gain estimates with the selection of inbred lines using the multiple-trait BLUP, especially when traits were negatively correlated. Viana et al. (2010) evaluated two traits in popcorn, which demonstrated a negative correlation (-0.39), and verified an increase in accuracies. Moreover, the authors confirmed that multiple-trait BLUP is efficient in the selection of half-sib families of annual crops. In this work, the values of mean accuracy of female parents, male parents, and families increased for all traits (Table 2), validating the efficiency of the use of multiple-trait BLUP in *Eucalyptus* breeding.

Despite high correlations in terms of ranking of families and the substantial agreement between the families selected for height, DBH, and volume by both methods, the selection gains using the multiple-trait BLUP were higher for all traits, regardless of the selection intensity. Therefore, these results, together with the higher accuracies achieved via multiple-trait BLUP, confirm the greater efficiency of this method when compared with the single-trait BLUP in the genetic selection of *Eucalyptus*. In future research, multiple-trait BLUP should be considered in association with genomic prediction analysis to improve accuracy and reduce the cycle time in the *Eucalyptus* breeding programs.

5 REFERENCES

Alves RS, Peixoto LA, Teodoro PE, Silva LA, Rodrigues EV, Resende MDV, Laviola BG, Bhering LL (2018) Selection of Jatropha curcas families based on temporal stability and adaptability of genetic values. Industrial Crops and Products 119:290-293.

Bauer AM, Léon J (2008) Multiple-trait breeding values for parental selection in self-pollinating crops. Theoretical and Applied Genetics 116:235-242.

Cohen J (1960) A coefficient of agreement for nominal scales. Educational and Psychological Measurement 20:37-46.

Corrêa TR, Picoli EAT, Souza GA, Condé SA, Silva NM, Lopes-Mattos KLB, Resende MDV, Zauza EAV, Oda S (2017) Phenotypic markers in early selection for tolerance to dieback in Eucalyptus. Industrial Crops and Products 107:130-138.

Costa RB, Resende MDV, Gonçalves PS, Silva MA (2002) Individual multivariate REML/BLUP in the presence of genotype x environment interaction in rubber tree (Hevea) breeding. Crop Breeding and Applied Biotechnology 2:131-140.

Costa RB, Silva JC, Skowronski L, Constantino M, Pistori H, Pinto JVC (2016) Genetic Divergence in Eucalyptus camaldulensis Progenies in the Savanna Biome in Mato Grosso, Brazil. PLoS ONE 11(9):e0163698.

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics. Longman Group Limited, Edinburgh.

Fonseca SM, Resende MDV, Alfenas AC, Guimarães LMS, Assis TF, Grattapaglia D (2010) Practical manual of genetic improvement of Eucalyptus. Editora UFV, Viçosa (in Portuguese).

Gilmour AR, Gogel BJ, Cullis BR, Welham SJ, Thompson R (2015) ASReml User Guide Release 4.1 Structural Specification. VSN International, Hemel Hempstead.

Harville DA, Carriquiry AL (1992) Classical and Bayesian prediction as applied to an unbalanced mixed linear model. Biometrics 48:987-1003.

Henderson CR (1975) Best linear unbiased estimation and prediction under a selection model. Biometrics 31(2):423-447.

Henderson CR, Quaas RL (1976) Multiple trait evaluation using relatives records. Journal of Animal Science 43(6):1188-1197.

Imai A, Kuniga T, Yoshioka T, Nonaka K, Mitani N, Fukamachi H, Hiehata N, Yamamoto M, Hayashi T (2016) Evaluation of the best linear unbiased prediction method for breeding values of fruit-quality traits in citrus. Tree Genetics & Genomes 12:119.

Kerr RJ (1998) Asymptotic rates of response from forest tree breeding strategies using best linear unbiased prediction. Theoretical and Applied Genetics 96:484-493.

Mrode RA (2014) Linear models for the prediction of animal breeding values. CAB International, Wallingford.

Patterson HD, Thompson R (1971) Recovery of inter-block information when block sizes are unequal. Biometrika 58:545-554.

Persson T, Andersson B (2004) Accuracy of single- and multiple-trait REML evaluation of data including non-random missing records. Silvae Genetica 53(3):135-139.

Piepho HP, Mohring J, Melchinger AE, Buchse A (2008) BLUP for phenotypic selection in plant breeding and variety testing. Euphytica 161:209-228.

Pollak EJ, Van der Werf J, Quaas RL (1984) Selection bias and multiple trait evaluation. Journal of Dairy Science 67:1590-1595.

R Development Core Team (2018) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Resende MDV (2015) Quantitative and Population Genetics. Suprema, Visconde do Rio Branco (in Portuguese).

Resende MDV (2016) Software Selegen-REML/BLUP: a useful tool for plant breeding. Crop Breeding and Applied Biotechnology 16:330-339.

Resende MDV, Ramalho MAP, Carneiro PCS, Carneiro JES, Batista LG, Gois IB (2016) Selection Index with Parents, Populations, Progenies, and Generations Effects in Autogamous Plant Breeding. Crop Science 56:530-546.

Resende MDV, Silva FF, Azevedo CF (2014) Mathematical, Biometric and Computational Statistics. Suprema, Visconde do Rio Branco (in Portuguese).

Schaeffer LR (2016) Random Regression Models. Available in http://animalbiosciences.uoguelph.ca/~lrs/BOOKS/rrmbook.pdf.

Silveira LCI, Brasileiro BP, Kist V, Weber H, Daros E, Peternelli LA, Barbosa MHP (2015) Selection strategy in families of energy cane based on biomass production and quality traits. Euphytica 204:443-455.

Viana JMS, Sobreira FM, Resende MDV, Faria VR (2010) Multi-trait BLUP in half-sib selection of annual crops. Plant Breeding 129:599-604.

Zhu Y, Wu S, Xu J, Lu Z, Li G, Hu Y, Yang X, Bush D (2017) Genetic parameters for growth traits and stem-straightness in Eucalyptus urophylla x E. camaldulensis hybrids from a reciprocal mating design. Euphytica 213(7):142.

6 SUPPLEMENTARY MATERIAL

Table S1. List of male parents.

Tree	Genetic material	Tree	Genetic material
1	E. grandis	33	Hyb. E. urophylla x (E. cam. x E. grandis)
2	E. grandis	34	Hyb. E. urophylla x (E. cam. x E. grandis)
3	E. grandis	35	Hyb. E. urophylla x (E. cam. x E. grandis)
4	E. grandis	36	Hyb. E. urophylla x (E. cam. x E. grandis)
5	E. saligna	37	Hyb. E. uro. x [E. uro. x (E. cam. x E. gran.)]
6	E. saligna	38	Hyb. E. uro. x [E. uro. x (E. cam. x E. gran.)]
7	E. urophylla	39	Hyb. E. uro. x [E. uro. x (E. cam. x E. gran.)]
8	E. urophylla	40	Hyb. E. camaldulensis x E. grandis
9	E. urophylla	41	Hyb. E. camaldulensis x E. grandis
10	E. robusta	42	Hyb. E. camaldulensis x E. grandis
11	E. resinifera	43	Hyb. E. camaldulensis x E. grandis
12	E. resinifera	44	Hyb. E. camaldulensis x E. grandis
13	E. resinifera	45	Hyb. E. camaldulensis x E. grandis
14	E. resinifera	46	Hyb. E. resinifera x E. grandis
15	E. resinifera	47	Hyb. E. resinifera x E. grandis
16	Hyb. spontaneous E. urophylla	48	Hyb. E. resinifera x E. grandis
17	Hyb. spontaneous E. urophylla	49	Hyb. E. resinifera x E. grandis
18	Hyb. spontaneous E. urophylla	50	Hyb. E. resinifera x E. grandis
19	Hyb. spontaneous E. urophylla	51	Hyb. E. resinifera x E. grandis
20	Hyb. spontaneous E. urophylla	52	Hyb. E. resinifera x E. grandis
21	Hyb. spontaneous E. urophylla	53	Hyb. E. urophylla x E. pellita
22	Hyb. spontaneous E. urophylla	54	Hyb. E. pellita x E. grandis
23	Hyb. spontaneous E. urophylla	55	Hyb. E. pellita x E. grandis
24	Hyb. spontaneous E. urophylla	56	Hyb. E. pellita x (E. gran. x E. teret.)
25	Hyb. spontaneous E. urophylla	57	Hyb. E. robusta x E. urophylla
26	Hyb. spontaneous E. urophylla	58	Hyb. E. robusta x E. grandis
27	Hyb. spontaneous E. urophylla	59	Hyb. E. robusta x E. grandis
28	Hyb. E. urophylla x E. grandis	60	Hyb. E. saligna x E. botryoides
29	Hyb. E. urophylla x E. grandis	61	Hyb. E. saligna x E. botryoides
30	Hyb. E. urophylla x E. grandis	62	Hyb. E. saligna x E. botryoides
31	Hyb. E. urophylla x E. grandis	63	Hyb. E. saligna x E. botryoides
32	Hyb. E. uro. x (E. cam. x E. grandis)	64	Hyb. E. saligna x E. grandis

CHAPTER 2

MULTIPLE-TRAIT BLUP FOR LONGITUDINAL DATA ANALYSIS IN JATROPHA CURCAS BREEDING

VIÇOSA - MINAS GERAIS

2020

ABSTRACT

Despite being a species with great potential for biodiesel production, little research has been done on the breeding of *Jatropha curcas*, mainly with respect to its yield across harvests. Thus, the present study was carried out to analyze longitudinal data via multiple-trait Best Linear Unbiased Prediction (BLUP) for the genetic improvement of *Jatropha curcas*. The experiment was set up as a randomized block design with two blocks and five plants per plot. The seed yield of 730 individuals of 73 half-sib families was evaluated over six years. Variance components and genetic parameters were estimated via Restricted Maximum Likelihood (REML). The Additive Index was used for ranking and selection purposes. Genetic correlations of low to moderate magnitude were observed between pairs of harvests. The Multiple-trait BLUP / Additive Index procedure allowed for the selection of superior families based on the predicted genetic values, considering all the harvests. Therefore, it can be efficiently applied in the breeding of *Jatropha curcas*.

Key words: Mixed model methodology; repeated measures; selection index; genetic selection

1 INTRODUCTION

The search for sustainable fuels has become a frequent topic in numerous world conferences. In this scenario, *Jatropha curcas* L. stands out as one of the main species from which biodiesel can be produced by virtue of its adaptability to different environments; its tolerance to drought, which enable its cultivation in low-fertile and non-irrigated soils (Laviola et al., 2017); and production life of more than 50 years (Nithiyanantham et al., 2012). Moreover, it has a higher capacity to transform oil into biodiesel compared with other species (Pu et al., 2011).

Despite being a species with great potential for biodiesel production, little research has been done on the breeding of *Jatropha curcas*. Recently, studies have been undertaken with a view to selecting genotypes using information of only one harvest (Junqueira et al., 2016); using several harvests independently (Bhering et al., 2013); or based on temporal stability and adaptability of genetic values (Alves et al., 2018). However, because *Jatropha curcas* is a perennial species that produces for many years, statistical methodologies that efficiently evaluate longitudinal data are necessary to allow for a more accurate genetic selection (Peixoto et al., 2016).

Longitudinal data analysis has a few peculiarities, as the many harvests are correlated with each other and because there may be heterogeneity of variances and covariances between them (Resende et al., 2014). The interest in analyses of this sort of data usually lies in the prediction of genetic values across all harvests as well as in the identification of a covariance structure over time. In plant breeding, longitudinal data associated with harvests are usually analyzed via repeatability models (Resende et al., 2014).

However, the multiple-trait Best Linear Unbiased Prediction (BLUP) — also known as the model with an unstructured covariance matrix — can be a very efficient alternative, because it uses all pieces of information simultaneously and treats repeated measures as different and correlated traits, considering their heritabilities and genetic and residual correlations (Mrode, 2014; Resende et al., 2014). This covariance structure is applied to all random factors of the statistical model such as genotype effects, plot effects, and residual effects (Gilmour et al., 2004). In this scenario, the present study proposes to analyze longitudinal data via multiple-trait Best Linear Unbiased Prediction (BLUP) for the genetic improvement of *Jatropha curcas*.

2 MATERIALS AND METHODS

2.1 EXPERIMENTAL DATA

The experiment was implemented in November 2008 as a randomized-block design with two blocks and five plants per plot that were arranged in rows spaced 4 m apart, with 2 m spacing between plants. The experiment was conducted in the experimental field of Embrapa Cerrados, located in Planaltina, DF - Brazil (15°35’30” S and 47°42’30” W; 1007 m asl). All management practices were based on Carels et al. (2012). The experiment consisted of an evaluation of the performance of 730 individuals of 73 half-sib families of *Jatropha curcas*. These individuals were evaluated for seed yield (kg plant^{-1}) in six harvests (2010 to 2015).

2.2 STATISTICAL ANALYSIS

The Restricted Maximum Likelihood/Best Linear Unbiased Prediction (REML/BLUP) procedure was adopted to estimate the variance components and predict the genetic values, in accordance with Patterson and Thompson (1971) and Henderson (1975).

The multiple-trait statistical model, associated with the evaluation of genotypes in a randomized-block design with several plants per plot, is given by the following equation:

$$y = Xr + Zf + Wp + e,$$

where y is the phenotypes vector; r is the vector of fixed effects of blocks added to the overall mean; f is the vector of random effects of families, $f \sim N(0, \sum_f \otimes I)$; p is the vector of random error effects between plots, $p \sim N(0, \sum_p \otimes I)$; e is the vector of random residual effects within plots, $e \sim N(0, \sum_e \otimes I)$; and X, Z, and W represent the incidence matrices for the mentioned effects. $\sum_f$ is a covariance matrix of random genetic effects of families; $\sum_p$ is a covariance matrix of random error effects between plots; $\sum_e$ is a covariance matrix of random residual effects within plots; I is an identity matrix of order appropriate to the respective random effect; and $\otimes$ denotes the Kronocker product. In addition, $\sum_f$, $\sum_p$, and $\sum_e$ are assumed unstructured covariance structures (US).

The additive genetic variance between family means (σ_a^2) was given by the following expression (Resende, 2015):

$$\sigma_a^2 = 4\sigma_f^2,$$

where σ_f^2 is the estimate of the genetic variance between family means.

The heritability between family means (h_f^2) was given by the following expression (Resende, 2015):

$$h_f^2 = \frac{\sigma_f^2}{\sigma_f^2 + \frac{1}{r}\sigma_{plot}^2 + \frac{1}{nr}\sigma_{res}^2},$$

where σ_{plot}^2 is the estimate of the variance between plots; σ_{res}^2 is the estimate of the residual variance within plots; *r* is the number of blocks; and *n* is the number of plants per plot.

The standard errors of the estimates of variance components and of the genetic parameters were obtained by the mixed model output and post-processing "pin" function of ASReml software.

Genetic correlations (Pearson's correlation) between pairs of harvests ($r_{g_{i,j}}$) were obtained based on the following expression:

$$r_{g_{i,j}} = \frac{\sigma_{f_{i,j}}}{\sqrt{\sigma_{f_i}^2 \sigma_{f_j}^2}},$$

where $\sigma_{f_{i,j}}$ is the genetic covariance estimated between the family means for the pair of harvests *i* and *j*; $\sigma_{f_i}^2$ is the genetic variance estimated between family means for harvest *i*; and $\sigma_{f_j}^2$ is the genetic variance estimated between family means for harvest j. The genetic covariances were obtained directly by the mixed model output.

The Additive Index (Resende, 2007), used to identify superior families to be selected in the *Jatropha curcas* breeding program, is given by the following expression:

$$AI_k = \sum_{h=1}^{6} w_h \frac{(u_h + f_{kh})}{\sigma_h},$$

where w_h is the weight assigned for harvest h; $u_h + f_{kh}$ is the overall mean for harvest h added to the predicted genetic effect of family k at harvest h; and σ_h is the standard deviation for $u_h + f_{kh}$. For the Additive Index, weights equal to u_h/u, where u is the overall mean, were assigned, and the selection direction was "higher". The gain predicted with the selection was obtained directly by the Additive Index output.

Statistical analyses were performed using ASReml 4.1 (Gilmour et al., 2015), Selegen REML/BLUP (Resende, 2016), and R (R Core Team, 2018) software.

3 RESULTS

The standard errors of the estimates of the variance components revealed that all the random effects of the statistical model are statistically significant (i.e., they differ from zero, with a confidence level of 66%) (Table 1). An increase was observed in the magnitude of genetic variance between family means and, consequently, in additive genetic variance between family means throughout the harvests, except for that of 2013. On the other hand, heritability between family means did not follow any pattern, ranging from 0.33 to 0.71 (2013 and 2014 harvests, respectively) (Table 1).

Table 1. Estimates of variance components and genetic parameters and their respective standard error, between parenthesis, for the seed yield trait, evaluated in 730 individuals of 73 half-sib families of *Jatropha curcas* in six harvests (2010-2015).

Harvest	σ_f^2 [a]	σ_a^2 [b]	σ_{plot}^2 [c]	σ_{res}^2 [d]	σ_{phen}^2 [e]	h_g^2 [f]
2010	0.0043 (0.0012)	0.0171 (0.0047)	0.0036 (0.0008)	0.0056 (0.0003)	0.0135 (0.0012)	0.65 (0.08)
2011	0.0099 (0.0055)	0.0397 (0.0218)	0.0284 (0.0057)	0.0314 (0.0018)	0.0697 (0.0056)	0.36 (0.15)
2012	0.1017 (0.0308)	0.4069 (0.1234)	0.1070 (0.0228)	0.1427 (0.0084)	0.3514 (0.0313)	0.60 (0.10)
2013	0.0693 (0.0428)	0.2772 (0.1712)	0.2512 (0.0457)	0.1372 (0.0081)	0.4577 (0.0429)	0.33 (0.16)
2014	0.2562 (0.0637)	1.0249 (0.2548)	0.1696 (0.0356)	0.2012 (0.0119)	0.6271 (0.0638)	0.71 (0.07)
2015	0.2922 (0.0772)	1.1689 (0.3089)	0.2235 (0.0496)	0.3099 (0.0186)	0.8257 (0.0785)	0.67 (0.08)

[a]: genetic variance between family means, [b]: additive genetic variance between family means, [c]: variance between plots, [d]: residual variance within plots, [e]: phenotypic variance, and [f]: heritability between family means.

Genetic correlations between pairs of harvests ranged from 0.11 (2011 and 2013) to 0.80 (2014 and 2015) (Figure 1), and 60% of these correlations were lower than 0.6. Overall, the highest genetic correlations were obtained for the last pairs of harvests.

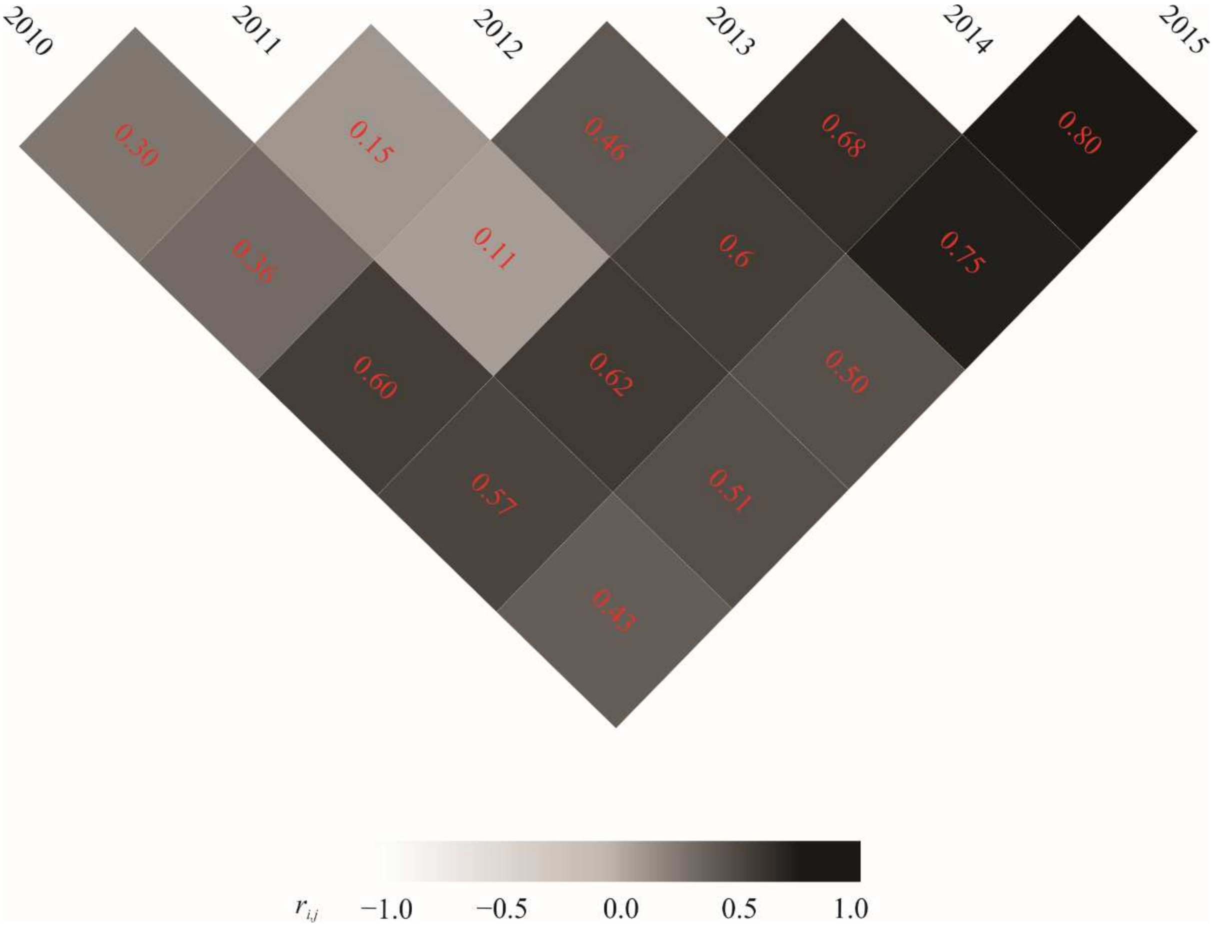

Figure 1. Genetic correlations between pairs of harvests for the seed yield trait, evaluated in 730 individuals of 73 half-sib families of *Jatropha curcas*, in six harvests (2010-2015).

Predicted genetic values in each harvest, Additive Index, and predicted selection gain across all harvests for the 73 half-sib families of *Jatropha curcas* are shown in the supplementary material - Table S1. By adopting a selection intensity of 27%, according to Additive Index, the 20 selected families were 41, 10, 15, 37, 36, 6, 54, 16, 9, 11, 72, 56, 18, 65, 34, 67, 73, 53, 70, and 39 (Figure 2), and the predicted selection gain across all harvests was 12%.

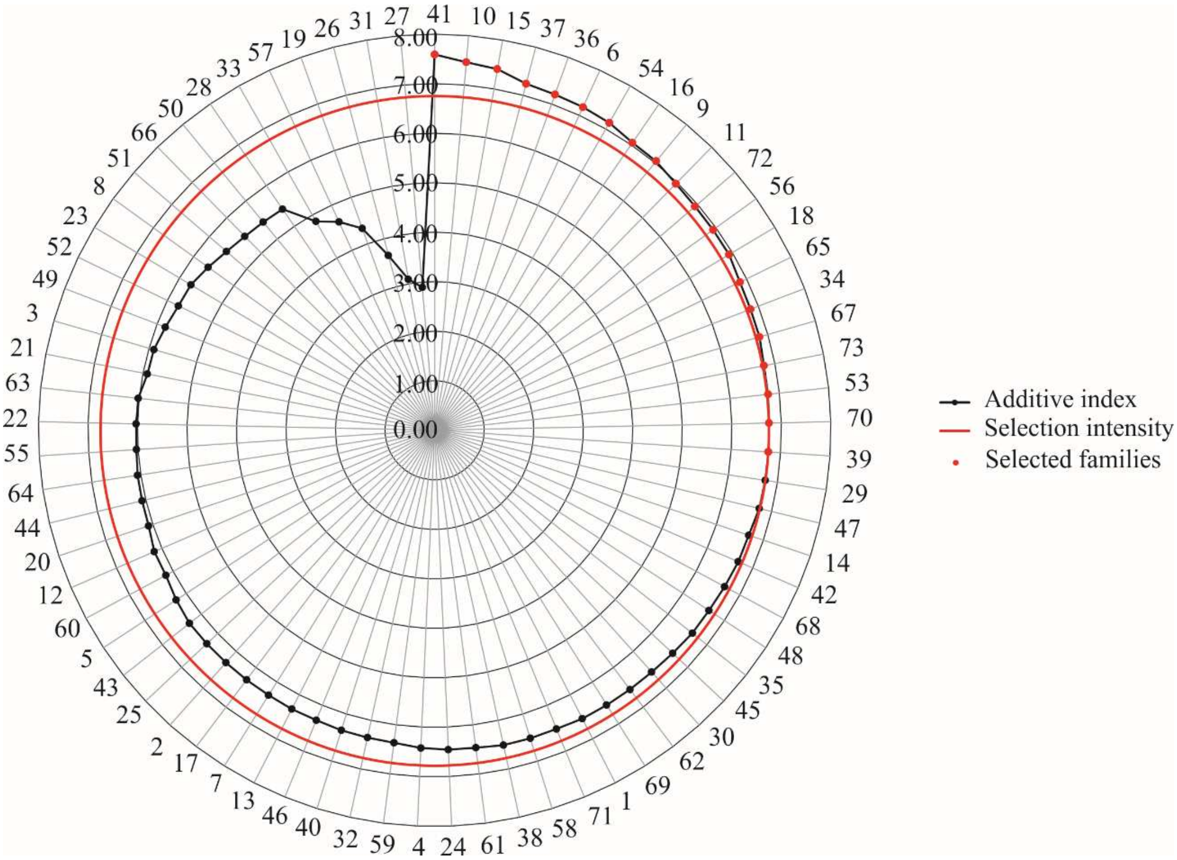

Figure 2. Ranking of *Jatropha curcas* families and selected families by the Additive Index.

4 DISCUSSION

BLUP assumes that variance components are known (Henderson, 1975). However, in practice, variance components are unknown and should thus be estimated in the most reliable way possible so that estimates can properly replace the parameters. In this case, the standard procedure for estimating variance components, under the approach of linear mixed models, is REML (Patterson and Thompson, 1971). Therefore, the genetic evaluation consists of using these methodologies together — the so-called REML/BLUP or mixed model methodology (Resende et al., 2014).

The main advantages of using the mixed model methodology are that it allows for the incorporation of kinship information; comparison of individuals or varieties over time and space; correction of environmental effects, simultaneous estimation of variance components and prediction of genetic values; and dealing with complex data structures. Moreover, the mixed model methodology can be applied to unbalanced data and non-orthogonal designs (Mrode, 2014; Resende, 2016).

In the identification and selection of superior families, genetic variance between families is characterized as one of the main components to quantify the breeding potential of the genotypes under study. Additionally, this parameter makes it possible to direct the selection strategy to be used, increasing the frequency of favorable alleles (Kerr, 1998).

The heritable proportion of total variability is named ‘heritability’ (Falconer and Mackay, 1996). This is a measure of genetic influence that informs that part of the population variation in a phenotype may be attributed to the variation in the genotype, allowing for estimates such as the genetic gain expected from selection (Allard, 1999). According to a scale proposed by Resende (2015), the heritabilities of the 2010, 2012, 2014, and 2015 harvests were characterized as high magnitude, whereas those of the 2011 and 2013 harvests were classified as moderate magnitude. These heritability estimates indicate a moderate to high correlation between the genotype and the phenotype, constituting a favorable scenario for the selection of families.

In the breeding of perennial plants, the number of measurements taken typically varies from three to six, since a higher number of harvests would compromise the efficiency of breeding programs per time unit (Resende, 2015). Laviola et al. (2013) stated that the efficiency of early selection in *Jatropha curcas* is small and that a minimum of 4 and 7 measurements are required to achieve reliability levels of 70 and 80%, respectively, in predicting the genetic values.

Repeatability models (the most simple and parsimonious) are very efficient when the genetic correlation between successive harvests is high (above 0.80) and are greatly advantageous because few parameters are to be estimated (Resende et al., 2014). However, in the current study, most genetic correlations between harvests were lower than 0.80, indicating that more complex models must be used. Thus, the multiple-trait BLUP is a more suitable strategy for the analysis of longitudinal data in *Jatropha curcas*. The predicted selection gain, with the recombination of the selected families, is of high magnitude and, therefore, confirms the efficiency of this procedure in the genetic improvement of *Jatropha curcas*.

Imai et al. (2016) applied multiple-trait BLUP in citrus and demonstrated the usefulness of the method in predicting genetic values when the information of kinship among genotypes is known and in the presence of unbalanced data. The authors recommended multiple-trait BLUP as a tool to select genotypes. Greater accuracy and efficiency of multiple-trait BLUP were also reported by Costa et al. (2002) in rubber tree (*Hevea brasiliensis*) breeding, by Kerr (1998) in the genetic selection of open and controlled cross-pollination cultivars in forestry breeding, and by Alves et al. (2018) in the genetic selection of *Eucalyptus*.

Few studies exist in the literature evaluating longitudinal data in *Jatropha curcas*. Thus, the results obtained in the present study using the multiple-trait BLUP / Additive Index procedure contribute to filling a gap in the breeding of *Jatropha curcas*. In addition, this procedure can be used for genetic selection in other species. In future research, multiple-trait BLUP should be considered in association with genomic prediction analysis to improve accuracy and reduce the cycle time in the genetic improvement programs of *Jatropha curcas*.

5 REFERENCES

Allard, R.W., 1999. Principles of plant breeding. Second edition. John Wiley & Sons, New York, USA.

Alves, R.S., Peixoto, L.A., Teodoro, P.E., Silva, L.A., Rodrigues, E.V., Resende, M.D.V., Laviola, B.G., Bhering, L.L., 2018. Selection of *Jatropha curcas* families based on temporal stability and adaptability of genetic values. IND CROP PROD 119, 290-293.

Alves, R.S., Rocha, J.R.A.S.C., Teodoro, P.E., Resende, M.D.V., Henriques, E.P., Silva, L.A., Carneiro, P.C.S., Bhering, L.L., 2018. Multiple-trait BLUP: a suitable strategy for genetic selection of *Eucalyptus*. TREE GENET GENOMES 14, 77.

Bhering, L.L., Barrera, C.F., Ortega, D., Laviola, B.G., Alves, A.A., Rosado, T.B., Cruz, C.D., 2013. Differential response of *Jatropha* genotypes to different selection methods indicates that combined selection is more suited than other methods for rapid improvement of the species. IND CROP PROD 41, 260-265.

Carels, N., Sujatha, M., Bahadur, B., 2012. *Jatropha*, Challenges for a New Energy Crop: Farming, Economics and Biofuel. Springer, New York, USA.

Costa, R.B., Resende, M.D.V., Gonçalves, P.D.S., Silva, M.D.A., 2002. Individual multivariate REML/BLUP in the presence of genotype x environment interaction in rubber tree (*Hevea*) breeding. CROP BREED APPL BIOT 131-140.

Falconer, D., Mackay, T., 1996. Introduction to Quantitative Genetics. Longman Scientific & Technical, Harlow, UK.

Gilmour A.R., Gogel, B.J., Cullis, B.R., Welham, S.J., Thompson, R., 2015. ASReml User Guide Release 4.1 Structural Specification. VSN International, Hemel Hempstead, UK.

Gilmour, A.R., Cullis, B.R., Welham, S.J., Gogel, B.J., Thompson, R., 2004. An efficient computing strategy for prediction in mixed linear models. COMPUT STAT DATA AN 44, 571-586.

Henderson, C.R., 1975. Best linear unbiased estimation and prediction under a selection model. BIOMETRICS 31 (2), 423-447.

Imai A., Kuniga, T., Yoshioka, T., Nonaka, K., Mitani, N., Fukamachi, H., Hiehata, N., Yamamoto, M., Hayashi, T., 2016. Evaluation of the best linear unbiased prediction method for breeding values of fruit-quality traits in citrus. TREE GENET GENOMES 12, 119.

Junqueira, V.S., Peixoto, L.A., Laviola, B.G., Bhering, L.L., Mendonça, S., Costa, T.S.A., Antoniassi, R., 2016. Bayesian Multiple-trait Analysis Reveals a Useful Tool to Increase Oil Concentration and to Decrease Toxicity in *Jatropha curcas* L. PLOS ONE 11, e0157038.

Kerr, R. J., 1998. Asymptotic rates of response from forest tree breeding strategies using best linear unbiased prediction. THEOR APPL GENET 96, 484-493.

Laviola, B.G., Oliveira, A.M.C., Bhering, L.L., Alves, A.A., Rocha, R.B., Gomes, B.E.L., Cruz, C.D., 2013. Estimates of repeatability coefficients and selection gains in *Jatropha* indicate that higher cumulative genetic gains can be obtained by relaxing the degree of certainty in predicting the best families. IND CROP PROD 51, 70-76.

Laviola, B.G., Rodrigues, E.V., Teodoro, P.E., Peixoto, L.A., Bhering, L.L., 2017. Biometric and biotechnology strategies in *Jatropha* genetic breeding for biodiesel production. RENEW SUST ENERG REV 76, 894-904.

Mrode, R.A., 2014. Linear models for the prediction of animal breeding values. Third edition, CAB International, Wallingford, Oxon, UK.

Nithiyanantham, S., Siddhuraju, P., Francis, G., 2012. Potential of *Jatropha curcas* as a biofuel, animal feed and health products. J AM OIL CHEM SOC 89, 961-972.

Patterson, H.D., Thompson, R., 1971. Recovery of inter-block information when block sizes are unequal. BIOMETRIKA 58, 545-554.

Peixoto, L.A., Laviola, B.G., Bhering, L.L., Mendonça, S., Costa, T.S.A., Antoniassi, R., 2016. Oil content increase and toxicity reduction in *Jatropha* seeds through family selection. IND CROP PROD 80, 70-76.

Pu, Y., Treasure, T., Gonzalez, R.W., Venditti, R., Jameel, H., 2011. Autohydrolysis pretreatment of mixed hardwoods to extract value prior to combustion. BIORESOURCES 6, 4856-4870.

R Core Team., 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Resende, M.D.V., 2007. SELEGEN REML/BLUP: sistema estatístico e seleção genética computadorizada via modelos lineares mistos. Embrapa Florestas, Colombo, Brasil.

Resende, M.D.V., 2015. Genética Quantitativa e de Populações. Suprema, Visconde do Rio Branco, Brasil.

Resende, M.D.V., 2016. Software Selegen-REML/BLUP: a useful tool for plant breeding. CROP BREED APPL BIOT 16: 330-339.

Resende, M.D.V., Silva, F.F., Azevedo, C.F., 2014. Estatística Matemática, Biométrica e Computacional. Suprema, Visconde do Rio Branco, Brasil.

6 SUPPLEMENTARY MATERIAL

Table S1. Predicted genetic values in each harvest (year), additive index (AI) and predicted selection gain (SG) (%) for the 73 half-sib families of *Jatropha curcas*.

Family	2010	2011	2012	2013	2014	2015	AI	SG%
41	0.3360	0.6387	1.7789	1.7795	3.1113	3.4764	7.59	21.12
10	0.2407	0.5464	2.2124	1.6789	3.0585	3.2043	7.46	20.12
15	0.3315	0.5871	2.1840	1.6747	2.9791	3.1128	7.40	19.48
37	0.2690	0.6653	1.6407	1.6223	3.0069	3.3932	7.24	18.49
36	0.3241	0.6670	1.6989	1.6167	2.9447	3.3095	7.20	17.79
6	0.2600	0.5723	1.7680	1.7479	2.8314	3.3098	7.18	17.27
54	0.2643	0.5882	1.8477	1.6518	2.7676	3.3199	7.14	16.80
16	0.2894	0.5166	1.9715	1.7464	2.8085	2.9820	7.05	16.26
9	0.2062	0.6343	1.8240	1.5016	3.0077	3.1436	7.04	15.83
11	0.2495	0.6406	1.7039	1.5362	2.7423	3.3133	6.97	15.36
72	0.3383	0.6180	1.6378	1.6528	2.7679	3.1392	6.93	14.94
56	0.2399	0.5111	1.7594	1.7390	2.7577	3.1131	6.93	14.58
18	0.2783	0.5517	1.8496	1.6297	2.8259	3.0096	6.92	14.27
65	0.2523	0.5930	1.8930	1.5100	2.4969	3.2644	6.85	13.92
34	0.2450	0.6752	1.6761	1.4687	2.7055	3.2206	6.83	13.60
67	0.2586	0.4809	1.9043	1.7100	2.4449	3.1500	6.82	13.31
73	0.2301	0.5609	1.5405	1.6804	2.6496	3.2223	6.78	13.01
53	0.2235	0.5355	1.7488	1.6160	2.4421	3.3008	6.77	12.74
70	0.2614	0.6986	1.0667	1.6595	2.7823	3.3798	6.76	12.48
39	0.2431	0.5805	1.5019	1.6806	2.6900	3.1558	6.75	12.24
29	0.1881	0.5609	1.8901	1.5084	2.5741	3.1358	6.75	12.03
47	0.2227	0.5908	1.8113	1.4906	2.6836	3.0671	6.74	11.83
14	0.2648	0.6513	1.7239	1.4255	2.7646	2.9731	6.69	11.61
42	0.2282	0.5086	1.9118	1.5871	2.5038	3.0138	6.68	11.41
68	0.2714	0.5686	1.5730	1.6377	2.4594	3.1899	6.66	11.20
48	0.1963	0.6195	1.4275	1.5696	2.6737	3.1787	6.63	10.99
35	0.2253	0.6266	1.6028	1.4933	2.5971	3.1298	6.63	10.80
45	0.2156	0.5112	1.7298	1.6024	2.4837	3.0626	6.59	10.60
30	0.3504	0.5901	1.5604	1.6091	2.7019	2.8052	6.56	10.39
62	0.2071	0.5784	1.7827	1.4594	2.3560	3.1710	6.56	10.20
69	0.2437	0.5481	1.6795	1.5496	2.5036	3.0352	6.55	10.02
1	0.2179	0.5823	1.5024	1.5984	2.6133	3.0286	6.54	9.85
71	0.3193	0.6178	1.4438	1.5918	2.5102	3.0322	6.52	9.67
58	0.2617	0.5623	1.4970	1.6035	2.5014	3.0781	6.52	9.50
38	0.2306	0.5723	1.5906	1.5258	2.4805	3.0993	6.51	9.34
61	0.2017	0.5407	1.7450	1.5108	2.2676	3.1407	6.46	9.17
24	0.3537	0.5502	1.5507	1.6988	2.5340	2.7490	6.46	9.01
4	0.1788	0.7358	1.4853	1.2547	2.6577	3.0808	6.42	8.84

Table S1. Continuation.

59	0.1950	0.6182	1.3109	1.4336	2.3975	3.3128	6.36	8.65
32	0.2964	0.6561	1.2298	1.5493	2.5536	2.9868	6.35	8.47
40	0.2542	0.5797	1.7256	1.4535	2.4485	2.8124	6.35	8.29
46	0.2130	0.5196	1.5875	1.5453	2.3781	2.9726	6.32	8.12
13	0.2258	0.5483	1.7551	1.4470	2.4866	2.7837	6.32	7.95
7	0.2323	0.5859	1.6116	1.4276	2.5885	2.7952	6.31	7.79
17	0.2523	0.6378	1.5433	1.3976	2.5622	2.8399	6.31	7.63
2	0.2035	0.7786	1.5040	1.1506	2.6385	2.9625	6.31	7.48
25	0.2998	0.5443	1.3599	1.6564	2.4990	2.8424	6.30	7.34
43	0.2475	0.4831	1.7301	1.5767	2.3973	2.7648	6.30	7.20
5	0.1628	0.5370	2.0484	1.3033	2.3166	2.7474	6.24	7.04
60	0.2215	0.4713	1.6503	1.5599	2.1742	2.8976	6.17	6.87
12	0.2702	0.5224	1.8365	1.4339	2.3280	2.6308	6.17	6.71
20	0.1669	0.5228	1.6768	1.3825	2.4128	2.7292	6.09	6.52
44	0.2455	0.5485	1.5853	1.4414	2.3198	2.7300	6.08	6.34
64	0.2182	0.4875	1.6924	1.4965	2.0077	2.9126	6.07	6.17
55	0.2980	0.5287	1.5733	1.4858	2.2051	2.7163	6.03	5.99
22	0.2296	0.6033	1.5070	1.3573	2.4034	2.7089	6.03	5.81
63	0.1700	0.5564	1.5027	1.3647	2.1137	3.0430	6.02	5.64
21	0.1860	0.4759	1.6332	1.4529	2.2936	2.5768	5.91	5.45
3	0.2032	0.6562	1.5296	1.2069	2.3730	2.6546	5.90	5.26
49	0.2655	0.5377	1.6184	1.3842	2.1487	2.5459	5.82	5.05
52	0.1703	0.5306	1.4896	1.3547	2.1028	2.7140	5.75	4.83
23	0.1714	0.5375	1.3176	1.3413	2.1650	2.8096	5.73	4.62
8	0.1740	0.6269	1.2302	1.2152	2.3270	2.6538	5.64	4.39
51	0.1680	0.4462	1.6823	1.3526	1.9896	2.4224	5.54	4.14
66	0.1870	0.5004	1.2594	1.3728	1.9722	2.6662	5.48	3.88
50	0.1956	0.5174	1.3630	1.3468	1.9764	2.5170	5.45	3.62
28	0.2106	0.5330	1.4015	1.3116	2.0079	2.4247	5.42	3.37
33	0.1643	0.4625	1.3543	1.1838	1.8942	2.0126	4.85	2.99
57	0.1416	0.5063	1.1262	1.0906	1.5264	2.3088	4.63	2.57
19	0.2091	0.4460	1.2023	1.1082	1.4484	1.8689	4.33	2.09
26	0.1310	0.3721	1.2220	0.9714	1.2008	1.3798	3.65	1.47
31	0.1092	0.3625	0.9694	0.8874	0.9019	1.2056	3.09	0.75
27	0.1209	0.3967	0.8088	0.7950	0.8991	1.1425	2.90	0.00
Mean	0.2319	0.5618	1.5939	1.4720	2.4129	2.8710	6.26	
Std deviation	0.0535	0.0757	0.2569	0.2001	0.4359	0.4571		