

## Evaluation of a new additive-dominance genomic model and implications for quantitative genetics and genomic selection

Taiana Lopes Rangel Miranda<sup>1</sup>, Marcos Deon Vilela de Resende<sup>2</sup>, Camila Ferreira Azevedo<sup>1</sup>, Andrei Caique Pires Nunes<sup>3</sup>, Elizabete Keiko Takahashi<sup>4</sup>, Guilherme Ferreira Simiqueli<sup>5</sup>, Fabyano Fonseca e Silva<sup>6</sup>, Rodrigo Silva Alves<sup>7\*</sup>

<sup>1</sup>Universidade Federal de Viçosa – Depto. de Estatística, Av. Peter Henry Rolfs s/n – 36570-977 – Viçosa, MG – Brasil.

<sup>2</sup>Embrapa Café, Parque Estação Biológica – 70770-901 – Brasília, DF – Brasil.

<sup>3</sup>Universidade Federal do Sul da Bahia/Centro de Formação em Ciências e Tecnologias Agroflorestais, Rod. BR-415, km 22 – 45613-204 – Itabuna, BA – Brasil.

<sup>4</sup>Celulose Nipo-Brasileira S.A, Rodovia BR 381, km 172 – 39740-000 – Belo Oriente, MG – Brasil.

<sup>5</sup>Universidade Federal de Viçosa – Depto. de Biologia Geral, Av. Peter Henry Rolfs s/n – 36570-900 – Viçosa, MG – Brasil.

<sup>6</sup>Universidade Federal de Viçosa – Depto. de Zootecnia, Av. Peter Henry Rolfs s/n – 36570-900 – Viçosa, MG – Brasil.

<sup>7</sup>Universidade Federal de Lavras/Instituto Nacional de Ciência e Tecnologia do Café, Campus Universitário – 37200-900 – Lavras, MG – Brasil.

\*Corresponding author <[rodrigo.s.alves@ufv.br](mailto:rodrigo.s.alves@ufv.br)>

**ABSTRACT:** The Fisher's infinitesimal model is traditionally used in quantitative genetics and genomic selection, and it attributes most genetic variance to additive variance. Recently, the dominance maximization model was proposed and it prioritizes the dominance variance based on alternative parameterizations. In this model, the additive effects at the locus level are introduced into the model after the dominance variance is maximized. In this study, the new parameterizations of additive and dominance effects on quantitative genetics and genomic selection were evaluated and compared with the parameterizations traditionally applied using the genomic best linear unbiased prediction method. As the parametric relative magnitude of the additive and dominance effects vary with allelic frequencies of populations, we considered different minor allele frequencies to compare the relative magnitudes. We also proposed and evaluated two indices that combine the additive and dominance variances estimated by both models. The dominance maximization model, along with the two indices, offers alternatives to improve the estimates of additive and dominance variances and their respective proportions and can be successfully used in genetic evaluation.

**Keywords:** REML, BLUP, genetic models, genetic selection, plant breeding

Edited by: Thomas Kumke

Received March 11, 2021

Accepted August 17, 2021

### Introduction

Genomic selection (GS), as proposed by Meuwissen et al. (2001), enables the identification of genetically superior individuals before their phenotypic data is collected and increases selection accuracy, accelerating and boosting the efficiency of genetic improvement (Simeão et al., 2021). GS emphasizes the simultaneous prediction of the genetic effects of thousands of genetic DNA markers dispersed throughout the genome of an individual in order to capture the effects of all loci and it explains all the genetic variance of a quantitative trait (Resende and Alves, 2020).

The Fisher's infinitesimal model (Fisher, 1918) attributes most genetic variance to additive variance and is traditionally used in quantitative genetics and GS. In the process of deriving biometric expressions, the additive variance is maximized while the dominance variance is the residue of the total genetic variance. In the dominance maximization model (Huang and Mackay, 2016), the dominance variance is prioritized using a parameterization in which the dominant homozygote and the heterozygote are weighed equally. The additive effects are introduced into the model after the dominance variance has been maximized.

The genomic best linear unbiased prediction (G-BLUP) is one of the methods commonly used in GS, which is suitable for predicting additive and dominance effects (Azevedo et al., 2015). G-BLUP predicts the genotypic effects of individuals using a mixed linear model. The additive-dominance G-BLUP captures the additive and dominance effects, allowing the effective genetic selection. This process maximizes the use of GS in animal and plant breeding (Azevedo et al., 2015).

Here, we analyzed and compared the quantitative genetics models proposed by Fisher (1918) and Huang and Mackay (2016) in terms of their effectiveness in estimating additive and dominance variances, heritabilities, and additive and dominance effects through the additive-dominance G-BLUP. We also proposed and evaluated two indices that combine the additive and dominance variances estimated by both models.

### Materials and Methods

#### Experimental data

In this study, we used data from the evaluation of a *Eucalyptus grandis* x *Eucalyptus urophylla* hybrid

population. The population involved 756 trees distributed in 37 outbred  $F_2$  full-sib families derived from the breeding of ten unrelated elite interspecific  $F_1$  hybrids. Trees were deployed in a field trial in a randomized complete block design with single-tree plots and 24-36 replications per family. At three years after planting, we evaluated the traits mean annual increment (MAI), basic density, and cellulose yield.

Single nucleotide polymorphism (SNP) markers were obtained using the Illumina Infinium (Gunderson et al., 2005), EuCHIP60K (Silva Junior et al., 2015). SNP markers were called from intensity files obtained through GENESEEK using GENOMESTUDIO 2011.1 following the standard procedures for genotyping and quality control with no manual editing of clusters. The average SNP call frequency across samples was > 90 % and the sample call rate across SNPs was > 95 %. SNP markers were then filtered by keeping SNPs with minor allele frequency (MAF) > 0.01, totaling 23,129 effective SNP markers.

### Genetic models

The standard approach for genetic evaluation is the mixed model methodology (Mrode and Thompson, 2014; Resende et al., 2014). This methodology needs individual genetic relationship matrices, which are A (pedigree based genetic relationships) for the BLUP (Henderson, 1975) and G (genomic based genetic relationships) for G-BLUP (Meuwissen et al., 2001). The original basic models for obtaining A and a modified G were raised by Fisher (1918) and Huang and Mackay (2016), respectively. In the Fisher's era, there was no G-BLUP but the basics for structuring A was already clearly stated and then inspired the development of G. Inspired by both studies mentioned, we used two) parameterizations of G in G-BLUP in this study.

### G-BLUP method

The additive-dominance genomic model for predicting the genotypic values of individuals was given as:

$$y = Xb + Zu_a + Zu_d + e \quad (1)$$

where:  $y$  is the vector of phenotypes ( $N \times 1$ , where  $N$  is the number of individuals);  $b$  is the vector of fixed effects ( $p \times 1$ , where  $p$  is the number of fixed effects) with incidence matrix  $X$  ( $N \times p$ );  $u_a$  and  $u_d$  are the vectors of additive and dominance effects ( $N \times 1$ ), respectively, with incidence matrix  $Z$  ( $N \times N$ ), assuming the existence of observations for all individuals; and  $e$  is the vector of residuals.

The variance structures were given  $u_a \sim N(0, G_a \sigma_a^2)$ , where  $\sigma_a^2$  is the additive variance, and  $G_a$  ( $N \times N$ ) is the genomic kinship matrix between individuals for additive effects;  $u_d \sim N(0, G_d \sigma_d^2)$ , where  $\sigma_d^2$  is the dominance

variance, and  $G_d$  ( $N \times N$ ) is the genomic kinship matrix between individuals for dominance effects; and  $e \sim N(0, I \sigma_e^2)$ , where  $\sigma_e^2$  is the residual variance, and  $I$  is an identity matrix ( $N \times N$ ).

The mixed model equations allowed predicting the additive ( $\hat{u}_a$ ) and dominance ( $\hat{u}_d$ ) effects through the additive-dominance G-BLUP method, as follows:

$$\begin{bmatrix} X'X & X'Z & X'Z \\ Z'X & Z'Z + G_a^{-1} \frac{\sigma_e^2}{\sigma_a^2} & Z'Z \\ Z'X & Z'Z & Z'Z + G_d^{-1} \frac{\sigma_e^2}{\sigma_d^2} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{u}_a \\ \hat{u}_d \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \\ Z'y \end{bmatrix} \quad (2)$$

where: the components of variance were estimated via restricted maximum likelihood (REML) method (Patterson and Thompson, 1971). Thus, the genotypic value for each individual was given as:

$$\widehat{GV} = \hat{u}_a + \hat{u}_d \quad (3)$$

### Traditional modeling

The traditional model uses the G-BLUP method with conventional parametrizations for the additive and dominance effects (Resende et al., 2014). The genomic relationship matrices for additive ( $G_a$ ) and dominance ( $G_d$ ) effects were given as:

$$G_a = \frac{WW'}{\sum_{i=1}^n 2p_i q_i} \quad \text{and} \quad (4)$$

$$G_d = \frac{SS'}{\sum_{i=1}^n (2p_i q_i)^2} \quad (5)$$

where:  $p_i$  and  $q_i$  are the allele frequencies,  $W$  is the additive incidence matrix, and  $S$  is the dominance incidence matrix.

For the additive incidence matrix ( $W$ ), the following parameterization was used:

$$W = \begin{cases} 2, & \text{genotype} = AA \\ 1, & \text{genotype} = Aa \\ 0, & \text{genotype} = aa \end{cases} \quad (6)$$

$$S = \begin{cases} 2(p - q), & \text{genotype} = AA \\ 2p, & \text{genotype} = Aa \\ 0, & \text{genotype} = aa \end{cases} \quad (7)$$

There are several parameterizations for the additive and dominance incidence matrices (Da et al., 2014; Resende et al., 2012; Resende et al., 2014; Varona et al., 2018; Vitezica et al., 2013). In this study, we followed the guidelines given by Huang and Mackay (2016).

**Alternative modeling**

The new model uses the G-BLUP method with alternative parameterizations for the additive and dominance effects and a novel way of calculating genomic relationship matrices (Huang and Mackay, 2016). Thus, the genomic relationship matrices for additive ( $G_a$ ) and dominance ( $G_d$ ) effects were given as:

$$G_a = \frac{WW'}{\sum_{i=1}^n \frac{2p_i^2 q_i}{1+q}} \text{ and} \tag{8}$$

$$G_d = \frac{SS'}{\sum_{i=1}^n \frac{4p_i q_i^2}{1+q}} \tag{9}$$

where:  $W$  and  $S$  contemplate the parameterizations proposed by Huang and Mackay (2016) for the additive and dominance effects. The additive incidence matrix ( $W$ ) was given as:

$$W = \begin{cases} \frac{-2q}{1+q}, & \text{genotype} = AA \\ \frac{1-q}{1+q}, & \text{genotype} = Aa \\ 0, & \text{genotype} = aa \end{cases} \tag{10}$$

For the dominance incidence matrix ( $S$ ), the dominant homozygote and heterozygote were coded with a value of 2, while the recessive homozygote was given a value of 0. Thus,  $S$  was defined as:

$$S = \begin{cases} 2, & \text{genotype} = AA \\ 2, & \text{genotype} = Aa \\ 0, & \text{genotype} = aa \end{cases} \tag{11}$$

The codes 2, 2 and 0 is just because, under a complete dominance model, the AA and Aa genotypes give rise to same effects on phenotypes. Table 1 shows the notations and definitions of the genetic variance components at the locus level used in this study.

**Indices**

Based on the results obtained from both models (traditional and alternative), two other genetic variance estimates were calculated: Index 1, which uses the average of the estimates obtained by traditional and alternative modeling and Index 2, which uses a weighting based on the standard deviation (measured from various samples of cross validation) of the estimates obtained by traditional and alternative modeling. Table 2 presents a summary of the four estimators used in this study.

**Cross validation**

The process of cross validation consisted of dividing the population into  $k$  groups. The individuals that

**Table 1** – Notations and definitions of the components of genetic variance at the locus level.

Notation	Variance component	Genotype coding (aa, Aa, AA)
$\sigma_a^2$	$2pq[a + d(q - p)]^2$	0, 1, 2
$\sigma_d^2$	$(2pqd)^2$	0, 2p, 2(p - q)
$\sigma_a'^2$	$\frac{2p^2q}{1+q}(a - d)^2$	0, $\frac{1-q}{1+q}, \frac{-2q}{1+q}$
$\sigma_d'^2$	$\frac{4pq^2}{1+q}(a + dq)^2$	0,2,2

$a$  and  $d$  = genotypic effects of homozygotes and heterozygotes, respectively;  $p$  and  $q$  = frequencies of alleles A and a, respectively;  $\sigma_a^2$  and  $\sigma_d^2$  = additive and dominance variances, respectively, based on the traditional model;  $\sigma_a'^2$  and  $\sigma_d'^2$  = additive and dominance variances, respectively, based on the alternative model.

**Table 2** – Description of the additive and dominance variances for the traditional and alternative models/indices

Modeling/Index	Additive variance	Dominance variance
Traditional	$\sigma_a^2$	$\sigma_d^2$
Alternative	$\sigma_a'^2$	$\sigma_d'^2$
Index 1	$\frac{\sigma_a^2 + \sigma_a'^2}{2}$	$\frac{\sigma_d^2 + \sigma_d'^2}{2}$
Index 2	$\frac{\sigma_a^2}{S_a} S_{a\text{mean}} + \frac{\sigma_a'^2}{S_a'} S_{a\text{mean}}$	$\frac{\sigma_d^2}{S_d} S_{d\text{mean}} + \frac{\sigma_d'^2}{S_d'} S_{d\text{mean}}$

$s_a$  and  $s_d$  = standard deviation of additive variance ( $\sigma_a^2$ ) and dominance variance ( $\sigma_d^2$ ) estimated in the validation groups by traditional modeling;  $s_a'$  and  $s_d'$  = standard deviation of additive variance ( $\sigma_a'^2$ ) and dominance variance ( $\sigma_d'^2$ ) estimated in the validation groups by alternative modeling; and  $s_{a\text{mean}}$  and  $s_{d\text{mean}}$  = average of  $s_a$  and  $s_a'$ , and of  $s_d$  and  $s_d'$ .

belonged to the  $k-1$  groups were used as the estimation population and the remaining group was used as the validation population. The statistical model was fitted into the estimation population. Subsequently, the genetic effects were calculated for individuals in the validation population using the effects of the markers already obtained in the estimation population. This process was repeated until all remaining groups were considered once as the validation population.

In this study, the number of groups also differed as the number of individuals was different for each evaluated trait. For MAI, we considered  $k = 12$  (63 individuals in each group), for basic density, we considered  $k = 16$  (47 individuals in each group), and for cellulose yield, we considered  $k = 7$  (107 individuals in each group). After cross validation, predictive capacities, heritabilities, and regression coefficients in each of the validation populations were obtained and the averages calculated.

**Comparison of the models**

The following efficiency measures for genomic predictions were calculated for each trait and used to compare both models: predictive capacity, regression coefficient, and heritability/coefficient of determination.

The genotypic predictive capacity ( $CP_g$ ), the additive predictive capacity ( $CP_a$ ), and the dominance predictive capacity ( $CP_d$ ) were calculated through the correlation between corrected phenotypic values and the predicted genotypic effects, predicted additive effects, and predicted dominance effects, respectively.

The regression coefficients between the corrected phenotypic values and the genotypic effects, additive effects, and dominance effects were given, respectively, as:

$$b_{y\hat{g}} = \frac{cov(y \cdot \hat{g})}{var(\hat{g})} \tag{12}$$

$$b_{y\hat{a}} = \frac{cov(y \cdot \hat{a})}{var(\hat{a})} \text{ and} \tag{13}$$

$$b_{y\hat{d}} = \frac{cov(y \cdot \hat{d})}{var(\hat{d})} \tag{14}$$

The broad sense heritability, narrow-sense heritability, and coefficient of determination of dominance effects were given, respectively, as:

$$h_g^2 = \frac{\sigma_a^2 + \sigma_d^2}{\sigma_a^2 + \sigma_d^2 + \sigma_e^2} \tag{15}$$

$$h_a^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_d^2 + \sigma_e^2} \text{ and} \tag{16}$$

$$c_d^2 = \frac{\sigma_d^2}{\sigma_a^2 + \sigma_d^2 + \sigma_e^2} \tag{17}$$

The parametric relative magnitudes of the additive and dominance effects vary with populations allelic frequencies (Le Roy, 1960). Therefore, we considered different MAF to compare the relative magnitudes.

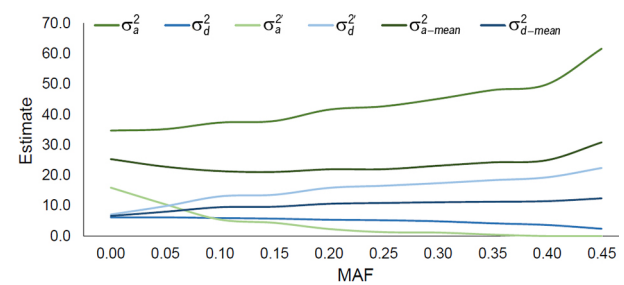
### Software

The statistical analyses were performed using the sommer package (Covarrubias-Pazarán, 2016) in R (R Development Core Team, version 3.4.0).

### Results

Table 3 shows the narrow-sense heritabilities, coefficients of determination of dominance effects, broad sense heritabilities, predictive capacities, and regression coefficients for the traits MAI, basic density, and cellulose yield considering the traditional and alternative models. In addition, the estimates of genetic variances for the traits MAI, basic density, and cellulose yield considering the traditional and alternative models and different MAF are shown in Figures 1, 2, and 3, respectively.

Basic density had the highest estimate of narrow-sense heritability when the traditional model was used (Table 3). This result was the opposite when the

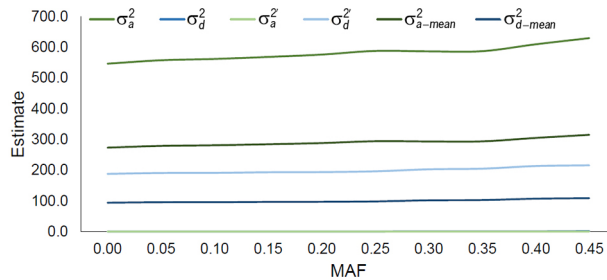


**Figure 1** – Estimates of genetic variances ( $\sigma_a^2$  = additive and  $\sigma_d^2$  = dominance, by the traditional model;  $\sigma_a^2$  = additive and  $\sigma_d^2$  = dominance, by the alternative model;  $\sigma_{a-mean}^2$  = mean additive and  $\sigma_{d-mean}^2$  = mean dominance) for mean annual increment (MAI) trait as a function of minor allele frequencies (MAF).

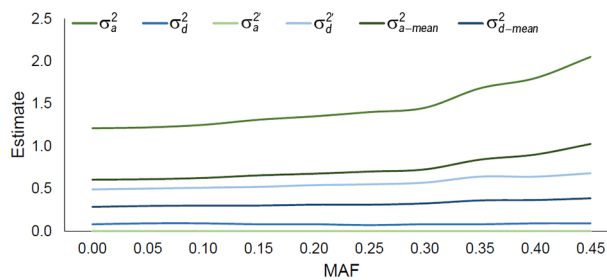
**Table 3** – Heritability/coefficient of determination estimates, predictive capacities, and regression coefficients according to the models for additive and dominance effects, minor allele frequency equal to one percent.

Trait	Modeling	$h_a^2$	$c_d^2$	$h_g^2$	$CP_g$	$CP_a$	$CP_d$	$b_{y\hat{a}}$	$b_{y\hat{d}}$	$b_{y\hat{g}}$
MAI	Traditional	0.47	0.08	0.55	0.58	0.56	0.04	0.92	0.38	0.97
	Alternative	0.28	0.13	0.41	0.57	0.56	0.57	2.52	1.84	0.97
	Index 1	0.38	0.11	0.48	0.58	0.56	0.31	1.72	1.11	0.97
	Index 2	0.48	0.11	0.59	-	-	-	-	-	-
Basic density	Traditional	0.56	0.00	0.56	0.62	0.62	-	1.00	-	1.00
	Alternative	0.00	0.31	0.31	0.58	-	0.58	-	0.99	0.99
	Index 1	0.28	0.16	0.44	0.60	0.62	0.58	1.00	0.99	1.00
	Index 2	0.22	0.09	0.31	-	-	-	-	-	-
Cellulose yield	Traditional	0.42	0.03	0.45	0.55	0.54	0.18	1.07	3.86	1.05
	Alternative	0.00	0.24	0.25	0.54	-	0.54	-	1.06	1.06
	Index 1	0.21	0.14	0.35	0.55	0.54	0.36	1.07	2.46	1.00
	Index 2	0.16	0.13	0.29	-	-	-	-	-	-

$h_a^2$  = narrow sense heritability;  $c_d^2$  = coefficient of determination of dominance effects;  $h_g^2$  = broad sense heritability;  $CP_g$  = genotypic predictive capacity;  $CP_a$  = additive predictive capacity;  $CP_d$  = dominance predictive capacity;  $b_{y\hat{a}}$  = regression coefficient between the corrected phenotypic values and additive effects;  $b_{y\hat{d}}$  = regression coefficient between the corrected phenotypic values and dominance effects;  $b_{y\hat{g}}$  = regression coefficient between corrected phenotypic values and genotypic effects.



**Figure 2** – Estimates of genetic variances ( $\sigma_a^2$  = additive and  $\sigma_d^2$  = dominance, by the traditional model;  $\sigma_a^2$  = additive and  $\sigma_d^2$  = dominance, by the alternative model;  $\sigma_{a\text{-mean}}^2$  = mean additive and  $\sigma_{d\text{-mean}}^2$  = mean dominance) for basic density trait as a function of minor allele frequencies (MAF).



**Figure 3** – Estimates of genetic variances ( $\sigma_a^2$  = additive and  $\sigma_d^2$  = dominance, by the traditional model;  $\sigma_a^2$  = additive and  $\sigma_d^2$  = dominance, by the alternative model;  $\sigma_{a\text{-mean}}^2$  = mean additive and  $\sigma_{d\text{-mean}}^2$  = mean dominance) for cellulose yield trait as a function of minor allele frequencies (MAF).

alternative model was applied, that is, the narrow-sense heritability for basic density becomes the lowest estimate (Table 3). Besides, the coefficients of determination of dominance effects estimated by the alternative model increased considerably compared to the effects estimated by the traditional model (Table 3).

The genotypic predictive capacity values were satisfactory for the traditional and alternative models and for all traits, ranging from 0.54 to 0.62, with slightly higher values for the traditional modeling. The same occurred for the regression coefficients, which ranged from 0.97 to 1.06.

The alternative model was efficient only for MAI trait (Table 3). For this trait, Indices 1 and 2 reveal that the practical estimates for narrow-sense heritabilities, coefficients of determination of dominance effects, and broad sense heritabilities were 0.38, 0.11, and 0.48; and 0.48, 0.11, and 0.59, respectively. On the other hand, in the traditional model, these estimates were 0.47, 0.08, and 0.55, respectively (Table 3).

Figures 1, 2, and 3 show an increase in dominance variance when estimated by the alternative modeling. More markers are eliminated with an increase in MAF. Therefore, in the traditional model, the estimates of additive variance ( $\sigma_a^2$ ) increase while the estimates of

dominance variance ( $\sigma_d^2$ ) decrease. The opposite occurs with the alternative model in which the higher the MAF, the higher the estimates of dominance variance ( $\sigma_d^2$ ) and the lower the estimates of additive variance ( $\sigma_a^2$ ). These results were observed for all traits (Figures 1, 2, and 3).

Figures 2 and 3 show high estimates of additive variance and low estimates of dominance variance in the model where additive variance ( $\sigma_a^2$ ) was maximized. In the model where dominance variance ( $\sigma_d^2$ ) was maximized, the estimates of additive variance ( $\sigma_a^2$ ) dropped to zero. These results show that the alternative model may not be appropriate for cases in which the dominance effects is irrelevant, as expected theoretically (Resende, 2015).

## Discussion

Although the traditional and alternative models are conceptually different, we can compare the estimated genetic variances/parameters between each other directly because of the same projections associated to the different maximizations imposed on the two objective functions. This ensures that the data (vector  $y$ ) is the same after the two different projections. The models scaled the estimated variance components in a way (considering the scale and the precision of the estimates) to avoid compatibility problems between these genetic variances. Then, interpretation of the meaning of these combining genetic variances is direct.

The traditional model for additive and dominance effects is driven by the fact that additive variance explains most of the genetic variance even under the dominance variance. In this model, both additive and dominance effects contribute to additive variance (Huang and Mackay, 2016). Thus, the dominance variance has little inference in the respective effects of dominance, because the additive variance is maximized first and the dominance variance is the residue of the total genetic variance.

Additive variance ( $\sigma_a^2$ ) is the sum of Type I squares of the regression of the genotypic effects in the number of copies of alleles, while dominance variance ( $\sigma_d^2$ ) is the residual genetic variance. Thus, priority is given to the additive component to explain the genetic variance. If priority is given to the dominance component using the alternative modeling in which heterozygote (Aa) and homozygote (AA) are coded identically (with code 2), an alternative dominance variance is defined,  $\sigma_d^2$  (Huang and Mackay, 2016). In this case, the dominance effects can explain the genetic variance first, while the additive variance ( $\sigma_a^2$ ) only enters the model after dominance variance ( $\sigma_d^2$ ) has been maximized. Thus, additive variance ( $\sigma_a^2$ ) becomes the residual genetic variance.

According to Resende (2015), one way to verify the relative importance of additive and dominance effects is to compare  $(\sigma_a^2 + \sigma_d^2)/2$  with  $(\sigma_a^2 + \sigma_a^2)/2$  with the greater value indicating which genetic effect is

more important. The model  $\sigma_a^2 + \sigma_d^2$  aims to maximize additive variance ( $\sigma_a^2$ ) and minimize dominance variance ( $\sigma_d^2$ ), while the model  $\sigma_a'^2 + \sigma_d'^2$  aims to maximize dominance variance ( $\sigma_d'^2$ ) and minimize additive variance ( $\sigma_a'^2$ ).

Inferences on population structure can be strongly influenced by the choice of MAF threshold (Linck and Battey, 2019). The results obtained here show that with an increase in MAF, the estimates of additive variance ( $\sigma_a^2$ ) in the traditional model also increase, while the estimates of dominance variance ( $\sigma_d^2$ ) decrease. However, higher estimates of dominance variance ( $\sigma_d^2$ ) and lower estimates of additive variance ( $\sigma_a^2$ ) are expected theoretically, since the dominance variance rises in relation to additive variance as the allelic frequency in the population increases (Resende, 2015). Therefore, dominance effects tend to be more important for breeding and selection in improved populations, which have higher frequencies of favorable alleles (Resende, 2002).

The alternative model is only appropriate when there is some degree of dominance acting on the genetic control of the trait (Resende, 2015). This occurs for MAI, but not for wood density and cellulose yields, which do not present allelic dominance (Rezende et al., 2014), explaining the zero values for additive variance obtained for wood density and cellulose yields. Therefore, it is essential to verify through the literature or previous data analyses if the trait in question presents any degree of allelic dominance. In natural populations, the effectiveness of using one of these models depends on the genetic control of the trait.

Index 1 considers the average of the additive and dominance variances estimates and presented good results for all traits, indicating that it is a good option. However, Index 2 weighs the standard deviation (precision) of the estimates obtained by the traditional and alternative models and is the most statistically appropriate.

The alternative model evaluated here showed effectiveness in estimating the components of genetic variance and predicting genetic values, when there are effects of dominance on the trait. After 100 years of using the traditional infinitesimal model, new alternatives are becoming available to evaluate quantitative traits (Visscher and Goddard, 2019).

## Conclusion

The alternative model presented interesting results and was appropriate for cases where both additive and dominance effects are relevant in the genetic control of the trait. The choice for model to be adopted in practice should consider not only practical knowledge of the trait (expression or not of heterosis for the average components), but also the estimates provided by both models and indices evaluated here, which give a weight to the estimates generated by both models.

## Acknowledgments

We acknowledge the financial support from National Institute of Science and Technology of Coffee (INCT Café), Minas Gerais State Agency for Research and Development (FAPEMIG), Brazilian National Council for Scientific and Technological Development (CNPq) and Coordination for the Improvement of Higher Level Personnel (CAPES) - Finance Code 001.

## Authors' Contributions

**Conceptualization:** Resende, M.D.V. **Data acquisition:** Takahashi, E.K. **Data analysis:** Miranda, T.L.R.; Nunes, A.C.P.; Simiqueli, G.F. **Design of methodology:** Miranda, T.L.R.; Resende, M.D.V.; Azevedo, C.F.; Nunes, A.C.P.; Simiqueli, G.F.; Silva, F.F. **Writing and editing:** Miranda, T.L.R.; Resende, M.D.V.; Alves, R.S.

## References

- Azevedo, C.F.; Resende, M.D.V.; Silva, F.F.; Viana, J.M.S.; Valente, M.S.F.; Resende Junior, M.F.R.; Muñoz, P. 2015. Ridge, Lasso and Bayesian additive-dominance genomic models. *BMC Genetics* 16: 105.
- Covarrubias-Pazarán, G. 2016. Genome-assisted prediction of quantitative traits using the R package sommer. *PLoS One* 11: e0156744.
- Da, Y.; Wang, C.; Wang, S.; Hu, G. 2014. Mixed model methods for genomic prediction and variance component estimation of additive and dominance effects using SNP markers. *PLoS One* 9: e87666.
- Fisher, R.A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh* 52: 399-433.
- Gunderson, K.L.; Steemers, F.J.; Lee, G.; Mendoza, L.G.; Chee, M.S. 2005. A genome wide scalable SNP genotyping assay using microarray technology. *Nature Genetics* 37: 549-554.
- Henderson, C.R. 1975. Best linear unbiased estimation and prediction under a selection model. *Biometrics* 3: 423-447.
- Huang, W.; Mackay, T.F.C. 2016. The Genetic Architecture of Quantitative Traits Cannot Be Inferred from Variance Component Analysis. *PLoS Genetics* 12: e1006421.
- Le Roy, H.L. 1960. Statistical methods of population genetics: a floor plan for geneticists, agronomists and biomathematicians = Statistische Methoden der Populationsgenetik. Ein Grundriss für Genetiker, Agronomen und Biomathematiker. Birkhäuser, Basel, Switzerland (in German).
- Linck, E.; Battey, C.J. 2019. Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. *Molecular Ecology Resources* 19: 639-647.
- Meuwissen, T.H.E.; Hayes, B.J.; Goddard, M.E. 2001. Prediction of total genetic value using genome wide dense marker maps. *Genetics* 157: 1819-1829.
- Mrode, R.A.; Thompson, R. 2014. Linear Models for the Prediction of Animal Breeding Values. CABI, Wallingford, UK.

- Patterson, H.D.; Thompson, R. 1971. Recovery of inter-block information when block sizes are unequal. *Biometrika* 58: 545-554.
- Resende, M.D.V. 2002. Biometric genetics and statistics in perennial plant breeding = Genética biométrica e estatística no melhoramento de plantas perenes. Embrapa, Brasília, DF, Brazil (in Portuguese).
- Resende, M.D.V. 2015. Quantitative and Population Genetics = Genética Quantitativa e de Populações. Suprema, Visconde do Rio Branco, MG, Brazil (in Portuguese).
- Resende, M.D.V.; Alves, R.S. 2020. Linear, generalized, hierarchical, Bayesian and random regression mixed models in genetics/genomics in plant breeding. *Functional Plant Breeding Journal* 2: 1-31.
- Resende, M.D.V.; Silva, F.F.; Azevedo, C.F. 2014. Mathematical, Biometric and Computational Statistics = Estatística Matemática, Biométrica e Computacional. Suprema, Visconde do Rio Branco, MG, Brazil (in Portuguese).
- Resende, M.D.V.; Silva, F.F.; Lopes, P.S.; Azevedo, C.F. 2012. Genomic Wide Selection (GWS) via Mixed Models (REML/BLUP), Bayesian Inference (MCMC), Multivariate Random Regression and Spatial Statistics = Seleção Genômica Ampla (GWS) via Modelos Mistos (REML/BLUP), Inferência Bayesiana (MCMC), Regressão Aleatória Multivariada e Estatística Espacial. UFV, Viçosa, MG, Brazil (in Portuguese).
- Rezende, G.D.S.P.; Resende, M.D.V.; Assis, T.F. 2014. Eucalyptus breeding for clonal forestry. p. 393-424. In: Fenning, T., ed. Challenges and opportunities for the world's forests in the 21<sup>st</sup> century. Springer, Dordrecht, Netherlands.
- Silva Junior, O.B.; Faria, D.A.; Grattapaglia, D. 2015. A flexible multi-species genome-wide 60K SNP chip developed from pooled resequencing 240 *Eucalyptus* tree genomes across 12 species. *New Phytologist* 206: 1527-1540.
- Simeão, R.M.; Resende, M.D.V.; Alves, R.S.; Pessoa-Filho, M.; Azevedo, A.L.S.; Jones, C.S.; Pereira, J.F.; Machado, J.C. 2021. Genomic selection in tropical forage grasses: current status and future applications. *Frontiers in Plant Science* 12: e665195.
- Varona, L.; Legarra, A.; Toro, M.A.; Vitezica, Z.G. 2018. Non-additive effects in genomic selection. *Frontiers in Genetics* 9: 78.
- Visscher, P.M.; Goddard, M.E. 2019. From RA Fisher's 1918 paper to GWAS a century later. *Genetics* 211: 1125-1130.
- Vitezica, Z.G.; Varona, L.; Legarra, A. 2013. On the additive and dominant variance and covariance of individuals within the genomic selection scope. *Genetics* 195: 1223-1230.